

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	1144	94.9	245	2	JQ0337	allergen Der p 1 -	
2	997.5	82.7	319	2	A61500	allergen Der f I p	
3	982.5	81.5	211	2	S21864	probable cysteine	
4	389.5	32.3	94	2	S03380	major fecal allergen	
5	333.5	27.7	358	2	JC7787	carrot seed cysteine	
6	332	27.5	348	2	F84672	probable cysteine	
7	327	27.1	348	2	JN0633	caricain (EC 3.4.2.	
8	327	27.1	367	2	JN0634	caricain (EC 3.4.2.	
9	326.5	27.1	326	2	T09259	cathepsin L-like p	
10	323.5	26.8	339	2	JC2476	cathepsin K (EC 3.	
11	322.5	26.7	374	2	T03941	cysteine proteinas	
12	322	26.7	346	2	C86413	cysteine proteinas	
13	322	26.7	362	1	S12581	cysteine proteinas	
14	321	26.6	368	2	S47312	cysteine proteinas	
15	321	26.6	466	2	T06416	cysteine proteinas	
16	317.5	26.3	329	2	A59468	cathepsin K (EC 3.	
17	315	26.1	325	2	A94828	ervatamin B (EC 3.	
18	313	26.0	458	1	KHRZ0A	oryzain (EC 3.4.22	
19	312.5	25.9	342	2	S71773	cysteine proteinas	
20	311.5	25.8	312	2	A23705	cysteine proteinas	
21	311.5	25.8	329	2	S74227	cathepsin K (EC 3.	
22	311	25.8	221	2	A59041	cysteine proteinas	
23	311	25.8	322	2	S19649	cysteine proteinas	
24	310.5	25.7	380	1	TAGB	actinidain (EC 3.4	
25	310	25.7	454	2	JC4848	cysteine proteinas	
26	310	25.7	471	1	KHRZ0B	oryzain (EC 3.4.22	
27	309.5	25.7	326	2	S43991	cathepsin L-like p	
28	308.5	25.6	323	2	S19650	cysteine proteinas	
29	308	25.5	365	2	T06206	probable cysteine	

```
Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGISNYCQIY 143
Qy 121 PPVANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSNA 180
Db 144 PPVANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSNA 203
Qy 181 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVUIL 222
Db 204 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVUIL 245

RESULT 2
A61500
allergen Der f I precursor - house-dust mite (Dermatophagoides farinae)
C:Species: Dermatophagoides farinae
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 26-Aug-1999
C:Accession: A61500
R:Dilworth, R.J.; Chua, K.Y.; Thomas, W.R.
C:Lin. Exp. Allergy 21, 25-32, 1991
A:Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I.
A:Reference number: A61500; MUID:91215493; PMID:2021874
A:Accession: A61500
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-319 <DIL>
C:Superfamily: papain

Query Match 82.7%; Score 997.5; DB 2; Length 319;
Best Local Similarity 81.2%; Pred. No. 3.2e-82;
Matches 181; Conservative 18; Mismatches 23; Indels 1; Gaps 1;

Qy 1 TNACISNG-NAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDLAE 59
Db 97 TSACRINSVNPSELDSLRLVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDLAE 156
Qy 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGISNYCQI 119
Db 157 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGISNYCQI 216
Qy 120 YPPNANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSN 179
Db 217 YPPDVQKIREALQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSN 276
Qy 180 AQGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVUIL 222
Db 277 TGGDDYWIVRNSWDTNMGDNGYGYFOAGNNLMIMEEQPYVWIM 319

RESULT 3
S21864
probable cysteine proteinase (EC 3.4.22.-) - Euroglyphus maynei
N:Alternate names: allergen Eur m I
C:Species: Euroglyphus maynei
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S21864
R:Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.
submitted to the EMBL Data Library, June 1991
A:Reference number: S21864
A:Accession: S21864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <KEN>
A:Cross-references: UNIPROT:P25780; EMBL:X60073
C:Genetics:
A:Introns: 100/3; 155/2
C:Superfamily: papain
C:Keywords: cysteine proteinase; hydrolase

Query Match 81.5%; Score 982.5; DB 2; Length 211;
Best Local Similarity 83.8%; Pred. No. 4.3e-81;
Matches 177; Conservative 17; Mismatches 16; Indels 1; Gaps 1;
```

```
Qy 1 TNACISNG-NAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDLAE 59
Db 1 TVACISINSVLSLESLDSLRLVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDLAE 60
Qy 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGISNYCQI 119
Db 61 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGLKNYCQI 120
Qy 120 YPPNANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSN 179
Db 121 SPDDNKIRQALQTHTAVAVIIGIKDLNFAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGSN 180
Qy 180 AQGVDDYWIVRNSWDTNMGDNGYGYFAANIDL 210
Db 181 TGGDDYWIVRNSWDTNMGDNGYGYFAANINL 211

RESULT 4
S03380
major fecal allergen Der p I - house-dust mite (Dermatophagoides pteronyssinus) (fragment)
N:Alternate names: allergen Der p1
C:Species: Dermatophagoides pteronyssinus
C>Date: 05-Mar-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S03380
R:Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.
C:Protein Seq. Data Anal. 2, 17-21, 1989
A:Title: Structural studies on the allergen Der p1 from the house dust mite Dermatophagoides pteronyssinus.
A:Reference number: A31657; MUID:89098855; PMID:2911558
A:Accession: S03380
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-28; 29-43; 44-60; 61-76; 77-94 <SIM>
A:Cross-references: UNIPROT:Q7M431
C:Superfamily: papain

Query Match 32.3%; Score 389.5; DB 2; Length 94;
Best Local Similarity 48.6%; Pred. No. 4.2e-28;
Matches 88; Conservative 0; Mismatches 4; Indels 89; Gaps 4;

Qy 1 TNACISNGNAPAEIDLQRMRTVTPIR--MQGGCGSWAFSGVAATESAYLAHRNQSILDLA 58
Db 1 TNACISNGNAPAEIDLQRMRTVTPIRMQGGCGSXXAFSGVA----- 43
Qy 59 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVVAREQSCRRPNAQRFGISNYCQ 118
Db 44 -----GIEYIQHNGVQESY-----RFGISNYCQ 68
Qy 119 YPPNANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHNAVNIYGS 178
Db 69 IYPPNANK-----DNGYQPNYXAVNIYGYX 93
Qy 179 N 179
Db 94 N 94

RESULT 5
JC7787
carrot seed cysteine proteinase (EC 3.4.-.-), CSCP - carrot
C:Species: Daucus carota (carrot)
C>Date: 02-Apr-2002 #sequence_revision 02-Apr-2002 #text_change 09-Jul-2004
C:Accession: JC7787
R:Sakuta, C.; Oda, A.; Konishi, M.; Yamakawa, S.; Kamada, H.; Satoh, S.
Biosci. Biotechnol. Biochem. 65, 2243-2248, 2001
A:Title: Cysteine proteinase gene expression in the endosperm of germinating carrot seeds
A:Reference number: JC7787
A:Contents: Seeds
A:Accession: JC7787
A:Molecule type: mRNA
A:Residues: 1-358 <SAK>
A:Cross-references: UNIPROT:Q948S1; DBJ:AB057371; DDBJ:AB057372
C:Comment: This protein is concerned with programmed cell death, and is endosperm-specific
C:Genetics:
```

A;Gene: cscop  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;151,286,307/Active site: Cys, His, Asn #status predicted

Query Match 27.7%; Score 333.5; DB 2; Length 358;  
Best Local Similarity 34.8%; Pred. No. 2.3e-22; Indels 21; Gaps 7;  
Matches 77; Conservative 36; Mismatches 87;

QY 11 PABIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQELVDCASQ- 69  
Db 128 PSSVDWRKKGAVPIKQKRCGSCWAFSAIVGVGINKITNQLVLSLSEQLVDCESDQ 187

QY 70 GCHGDTIPRGIEYIQHN-GVQSSYYRYVAREQSC--RRPNAQRFGISNVCQIYPPNANK 126  
Db 188 GCNGGLMENALEFTKRSGGVTVTERVYPYRARDERCATKQNPVVKIDGHENVPENNEYA 247

QY 127 IREALAQTHSAIAVIGIKDLDAFRH--YDGRITTIQRDNGYQPNYHVNIVGV-SNAQGV 183  
Db 248 LAQAVANQPVSVDAIDAGSDMQFYREGVYTGECGTELD-----HGVAVVGYGATNDGT 300

QY 184 DYWIVRNSWDTNMGDNGYGYFAANID-----LWMIEEYP 217  
Db 301 KYWIVKNSWGTDMGERGYIWRVDRINAAEGICGIAEASYP 341

RESULT 6  
F84672  
probable cysteine proteinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: F84672  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavay, L.; Tallon, L.;  
eues, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84672  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-348 <STO>  
A;Cross-references: UNIPROT:Q9ZQH7; GB:AE002093; NID:g4314384; PIDN:AAD15594.1; GSPDB:GN  
C;Genetics:  
A;Map position: 2  
C;Superfamily: papain

Query Match 27.5%; Score 332; DB 2; Length 348;  
Best Local Similarity 36.2%; Pred. No. 3e-22;  
Matches 84; Conservative 32; Mismatches 86; Indels 30; Gaps 11;

QY 5 SINGNAPAEIDLROMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQELVD 64  
Db 126 SDNGES---MDWRQEGAVTPVKYQRCGCGCWAFSAVAAGITKTKGELVLSUSEQQLLD 182

QY 65 CASOH--GCHGDTIPRGIEY-IQHNQVQSSYYRYVAREQSCRPNA-----QRFGISNY 116  
Db 183 CDRDYNQCGRGKINSKAFYIKKQGITTEDNTPYQESQTCSSSTLSSFRATISGY 242

QY 117 CQIYPPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITTIQRDNGY--QPNYHVN 173  
Db 243 -ETVPMNN---EEALLQAVSQPVSIGIEGTGAFRHYSQGVF-----NGECGTDLHHAVT 294

QY 174 IVGYS-NAQGVVDYWIVRNSWDTNMGDNGYGYFAANID-----LWMIEEYP 217  
Db 295 IVGYGMSSEEGTKYVWVKNWSWGETWGENGYMRKRDVDAPQGMCGLATIATYP 346

RESULT 7  
JN0633  
Carica (EC 3.4.22.30) I precursor - papaya  
N;Alternate names: papaya proteinase omega I; peptidase II; proteinase A; p

C;Species: Carica papaya (papaya)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: JN0633; S01135; S11748; A43026; S35728  
R;Revell, D.F.; Cummings, N.J.; Baker, K.C.; Collins, M.E.; Taylor, M.A.J.; Sumner, I.G.  
Gene 127, 221-225, 1993  
A;Title: Nucleotide sequence and expression in Escherichia coli of cDNAs encoding papaya  
A;Reference number: JN0633; MUID:93273235; PMID:7684720  
A;Accession: JN0633  
A;Molecule type: mRNA  
A;Residues: 1-348 <REV>  
A;Cross-references: UNIPROT:P10056; EMBL:X66060; NID:g18097; PIDN:CAA46862.1; PID:g18098  
A;Experimental source: fruit and leaf  
R;Dubois, T.; Kleinschmidt, T.; Schnek, A.G.; Looze, Y.; Braunitzer, G.  
Biol. Chem. Hoppe-Seyler 369, 741-754, 1988  
A;Title: The thiol proteinases from the latex of Carica papaya L. II. The primary structure  
A;Reference number: S01135; MUID:89105320; PMID:3063283  
A;Accession: S01135  
A;Molecule type: protein  
A;Residues: 133-348 <DUB>  
R;Collins, M.E.; Revell, D.F.; Sumner, I.G.; Pickersgill, R.W.; Goodenough, P.W.  
submitted to the EMBL Data Library, February 1990  
A;Reference number: S11748  
A;Accession: S11748  
A;Molecule type: mRNA  
A;Residues: 237-348 <COL>  
A;Cross-references: EMBL:X51899; NID:g18095; PIDN:CAA36180.1; PID:gl360713  
C;Genetics:  
A;Gene: Pp-omega  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;1-14/Domain: signal sequence #status predicted <SIG>  
F;15-132/Domain: propeptide #status predicted <PRO>  
F;133-348/Product: caricain #status experimental <MAT>  
F;154-195,188-227,285-336/Disulfide bonds: #status predicted  
F;157,291,311/Active site: Cys, His, Asn #status predicted

Query Match 27.1%; Score 327; DB 2; Length 348;  
Best Local Similarity 37.2%; Pred. No. 8.4e-22;  
Matches 74; Conservative 34; Mismatches 77; Indels 14; Gaps 7;

QY 9 NAPAETDLROMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQELVDCASQ 68  
Db 132 NLPENVDWRKKGAVTPVRHGGCGSCWAFSAVAVTEGINKIRTKLVELSEQLVDCERR 191

QY 69 -HCHGDTIPRGIEYIQHNQVQSSYYRYVAREQSCRRPNAQRFG-----ISNYCQIYPP 122  
Db 192 SHGCKGQYPPYALEYVAKNGIHLRSKYPYKAKQGTCK--AKQVGGPIVKTSGVGRVQPN 248

QY 123 NANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITTIQRDNGYQPNYHVNIVGYSNAQ 182  
Db 249 NEGNLLNATAK--QPVSVVVESKG-RPFQLYKG-GIFEGPCGTKVD-HAVTAVGYGKSGG 303

QY 183 VDYWIVRNSWDTNMGDNGY 201  
Db 304 KGVILKNWSWGTANGKGY 322

RESULT 8  
JN0634  
Carica (EC 3.4.22.30) II precursor - papaya  
N;Alternate names: papaya proteinase omega II  
C;Species: Carica papaya (papaya)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: JN0634; A43027; S31823  
R;Revell, D.F.; Cummings, N.J.; Baker, K.C.; Collins, M.E.; Taylor, M.A.J.; Sumner, I.G.  
Gene 127, 221-225, 1993  
A;Title: Nucleotide sequence and expression in Escherichia coli of cDNAs encoding papaya  
A;Reference number: JN0633; MUID:93273235; PMID:7684720  
A;Accession: JN0634  
A;Molecule type: mRNA  
A;Residues: 1-367 <REV>  
A;Cross-references: UNIPROT:Q42673; EMBL:X69877; NID:g22660; PIDN:CAA49504.1; PID:g22661  
A;Experimental source: fruit and leaf

C:Genetics:  
A:Gene: Pp-omega  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:1-11/Domain: signal sequence #status predicted <SIG>  
F:12-132/Domain: propeptide #status predicted <PRO>  
F:133-367/Product: carlacin II #status predicted <MAT>  
F:154-195,188-227,285-336/Distal folds: #status predicted  
F:157,291,311/Active site: Cys, His, Asn #status predicted

Query Match 27.1%; Score 327; DB 2; Length 367;  
Best Local Similarity 37.2%; Pred. No. 9e-22;  
Matches 74; Conservative 34; Mismatches 77; Indels 14; Gaps 7;

QY 9 NAPAEDLRQMTVTPIRMQGGSCWAFSGVAATESAYLAHRNQSLLDLAEQLVDCASQ 68  
DB 132 NLPENVDWRKKAATVPRHOGSCGSCWAFSAVATVEGINKIRTKGLVELSEQLVDCERR 191  
QY 69 -HGCHGDTTIPRGIEYIOHNGVQESYRYVAREQSCRRPNAQRFQ-----ISNYCQIYPP 122  
DB 192 SHGCKGGYPPYALEYVAKNGIHLRSKYPYKAKQGTCTR---AKQVGGPIVKTSGVGRVQPN 248  
QY 123 NANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRTHIQRDNGYQPNYHVNIVGYSNAQG 182  
DB 249 NEGNLLNIAIAK--QPVSVVVESKG-RPFQLYKG-GIFEGPCGTQVD-HAVTAVGYKSGG 303  
QY 183 VDIWIVRNSWDTNWGNGY 201  
DB 304 KGYILKNISWGTAWGEKGY 322

RESULT 9  
T09259  
cathepsin L-like proteinase (EC 3.4.22.-) - liver fluke  
C:Species: Fasciola hepatica (liver fluke)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09259  
R:Heuser, V. T.; Dobbelaere, D. A. E.  
Mol. Biochem. Parasitol. 64, 11-23, 1994  
A:Title: Cloning of a protease gene family of Fasciola hepatica by polymerase chain reaction  
A:Reference number: Z16631; MUID:94359526; PMID:8078514  
A:Accession: T09259  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-326 <HEU>  
A:Cross-references: UNIPROT:Q24944; EMBL:Z22765; NID:G452257; PID:G452258  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:132,269,289/Active site: Cys, His, Asn #status predicted

Query Match 27.1%; Score 326.5; DB 2; Length 326;  
Best Local Similarity 34.4%; Pred. No. 8.7e-22;  
Matches 72; Conservative 31; Mismatches 83; Indels 23; Gaps 5;

QY 11 PAEDLRQMTVTPIRMQGGSCWAFSGVAATESAYLAHRNQSLLDLAEQLVDCASQ---S 67  
DB 109 PESIDRDRYYTEVKDQGGSCWAFSTTGAVGQFRKNERASFSQQLVDCRTRDFG 168  
QY 68 QHCHGDTTIPRGIEYIOHNGVQESYRYVAREQSCRRPNAQRFQ-ISNYCQIYPPNANK 126  
DB 169 NYGCGGGYMEAYELKHNGLETESYYPYQAVEGPCQYDGLAYAKVTGYTYVHSGDEIE 228  
QY 127 IREALAQTHSAIAVIIGIKDLDAFR-HYDGRTHIQRDNGYQPN-----YHAVNIVGYS 178  
DB 229 LKN-----LVGTEDLPAVALDADSDFMVQSGIYQSGTCLPDLRTHAVLAVGYG 277  
QY 179 NAQGVVDYIWRNSWDTNWGNGYGYFAAN 207  
DB 278 SQDGTDYIWKNSWGTWGEDGYIRFARN 306

RESULT 10  
JC2476

cathepsin K (EC 3.4.22.-) precursor - human  
N:Alternate names: cathepsin O2  
C:Species: Homo sapiens (man)  
C:Date: 21-Feb-1995 #sequence\_revision 05-Apr-1995 #text\_change 09-Jul-2004  
C:Accession: JC2476; S55763; S68459; I38752; S48830  
R:Inaoka, T.; Bilbe, G.; Ishibaashi, O.; Tezuka, K.; Kumegawa, M.; Kokubo, T.  
Biochem. Biophys. Res. Commun. 206, 89-96, 1995  
A:Title: Molecular cloning of human cDNA for cathepsin K: Novel cysteine proteinase predicted  
A:Reference number: JC2476; MUID:95118380; PMID:7818555  
A:Accession: JC2476  
A:Molecule type: mRNA  
A:Residues: 1-329 <INA>  
A:Cross-references: UNIPROT:P43235; EMBL:X82153; NID:G562756; PIDN:CAA57649.1; PID:G562756  
R:Broemme, D.; Okamoto, K.  
Biol. Chem. Hoppe-Seyler 376, 379-384, 1995  
A:Title: Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastoma  
A:Reference number: S55763; MUID:96082523; PMID:7576232  
A:Accession: S55763  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-329 <BRO>  
A:Cross-references: GB:S79895; NID:G1195555; PIDN:AAB35521.1; PID:G1195556  
R:Shi, G. P.; Chapman, H. A.; Bhairi, S. M.; Deleeuw, C.; Reddy, V. Y.; Weiss, S. J.  
FEBS Lett. 357, 129-134, 1995  
A:Title: Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of  
A:Reference number: I38752; MUID:95104457; PMID:7805878  
A:Accession: S68459  
A:Molecule type: mRNA  
A:Residues: 1-329 <SHI>  
A:Cross-references: EMBL:U13665; NID:G606922; PIDN:AAA65233.1; PID:G606923  
A:Experimental source: tissue-type blood  
C:Genetics:  
A:Gene: GDB:CTSK; PYCD; CTSO1; CTSO2; PKND; CTSO  
A:Cross-references: GDB:453910; OMIM:265800; OMIM:601105  
A:Map position: 1q21-q21  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; glycoprotein; hydrolase  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-114/Domain: propeptide #status predicted <PRO>  
F:115-329/Product: cathepsin K #status predicted <MAT>  
F:103,161,213/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:139,276,296/Active site: Cys, His, Asn #status predicted

Query Match 26.8%; Score 323.5; DB 2; Length 329;  
Best Local Similarity 36.0%; Pred. No. 1.6e-21;  
Matches 73; Conservative 30; Mismatches 95; Indels 5; Gaps 5;

QY 8 GNAPAEIDLQRMRTVTPIRMQGGSCWAFSGVAATESAYLAHRNQSLLDLAEQLVDCAS 67  
DB 113 GRAPDSVDYRKGYVTVPKNQGGSCWAFSGVALEGGQLKKTKGLKLLNLSQNLVDCVS 172  
QY 68 QH-GCHGDTTIPRGIEYIOHNGVQESYRYVAREQSCRRPNAQRFQISNYCQIYPPNA 124  
DB 173 ENDGCGGGYMTWAFQVQKNRGIDSEDAIYYPYQGESCMYNTPTGKAACRCGYREIPEGNE 232  
QY 125 NKIREALAQTHSAIAVIIGIKDLDAFRHYDGRTHIQRDNGYQPNYHVNIVGYSNAQGV 184  
DB 233 KALKAVARV-GPVSVAID-ASLTGFSQYKSGYVDESCNLDNLNHLAVLAVGYGKGNK 290  
QY 185 YWIVRNSWDTNWGNGYGYFAAN 207  
DB 291 HWIKNSWGNKNGYILMARN 313

RESULT 11  
T03941  
cysteine proteinase (EC 3.4.22.-) precursor - common tobacco  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T03941  
R:Becker, C.; Muentz, K.  
submitted to the EMBL Data Library, September 1997  
A:Description: cDNA cloning of a CPRI-homologous proteinase from germinating tobacco seed



A;Reference number: Z15148  
A;Accession: F03941  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-374 <BEC>  
A;Cross-references: UNIPROT:O24137; EMBL:Z99173; PIDN:CABL6317.1  
A;Experimental source: clone TCPR1  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-374/Product: cysteine proteinase #status predicted <MAT>  
F;164,299,319/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322.5; DB 2; Length 374;  
Best Local Similarity 36.8%; Pred. No. 2.3e-21;  
Matches 81; Conservative 28; Mismatches 92; Indels 19; Gaps 9;

QY 11 PAETIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDC--ASQ 68  
Db 141 PHSVDMRRKGAAPIKQGGCGSCWAFSTVAAVEGINQIVTGEMITLSEQELVDCDRVQN 200  
QY 69 HGCHGDTIPRGIEYIQNGVQV--ESYRYVAREQSCR--RPNAQRFGISNYCQIYPPNAN 125  
Db 201 SGCNGGLMDYAFEFIIISNGGMDTEKHYPRGVGRCDPVRKNYKVSIDGYDV-PRNER 259  
QY 126 KIREALQTHSAIAVIIGIKDLDAFRHYDGRITQIRONGQPNVHNVISNAQGVYD 185  
Db 260 ALOKAVA--HQPVCAIEASG-RAFQLYSS-GVFTGCGBEVD-HGVVVVVGSGEDGVY 314  
QY 186 WIVRNSDWTWNGDNGYCPAANI-----DLMMIEEYP 217  
Db 315 WIVRNSGKTWNGYGVYMERNVKSHLGKGINTEASYP 354

RESULT 12  
C86413  
cysteine proteinase (EC 3.4.22.-) [similarity] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: C86413  
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: C86413  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-346 <STO>  
A;Cross-references: UNIPROT:Q9LP39; GB:AB05172; NID:G9502427; PIDN:AAF88126.1; GSPDB:GN  
C;Genetics:  
A;Map position: 1  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;154,289,310/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322; DB 2; Length 346;  
Best Local Similarity 32.3%; Pred. No. 2.4e-21;  
Matches 74; Conservative 33; Mismatches 86; Indels 36; Gaps 7;

QY 1 TNACSIGNPAEII-----DLRQMTVTPIRMQGGCGSCWAFSGV 40  
Db 101 TGLRGVNVTPFVFNVTETKPAWNTVSDVLGTNKNRNEGAVTPVKSQGGCGWAFSAI 160  
QY 41 AATESAYLAHRNQSLLDAEQELVDCSQ--HGCHGDTIPRGIEY-IQNGVQVESYRYV 97  
Db 161 AAVEGLTKIARGNLISLSEQQLDCTREQNNGCKGGTFVNAFNIIKRGISSENEYPYQ 220

QY 98 AREQSCRRPNAQRFGISNYCQIYPPNANKIREALQTHSAIAVIIGIKDLDA-----PRHY 153  
Db 221 VKEGPKRSNARPAAILIRGPENVPNNERALLAVSRQPVAVAI-----DASEAGFVHY 273  
QY 154 DGRITQIRONGQPNVHNVISNAQGVYDVIWRNSDWTWNGDNGY 201  
Db 274 SGGVYNARNCGTSVN-HAVTLVGVTGSPGKMYLAKNSWGTWNGY 321

## RESULT 13

S12581  
cysteine proteinase (EC 3.4.22.-) precursor - black gram  
N;Alternate names: cysteine endopeptidase; sulphydryl endopeptidase  
C;Species: Vigna mungo (black gram)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S12581; S05497; S20213; S48684  
R;Akasofu, H.; Yamauchi, D.; Minamikawa, T.  
Nucleic Acids Res. 18, 1892, 1990  
A;Title: Nucleotide sequence of the gene for the Vigna mungo sulphydryl-endopeptidase (S)  
A;Reference number: S12581; MUID:90245586; PMID:2336365  
A;Accession: S12581  
A;Molecule type: DNA  
A;Residues: 1-362 <AKA1>  
A;Cross-references: UNIPROT:P12412; EMBL:X51900; NID:922065; PIDN:CAA36181.1; PID:922066  
R;Akasofu, H.; Yamauchi, D.; Mitsuhashi, W.; Minamikawa, T.  
Nucleic Acids Res. 17, 6733, 1989  
A;Title: Nucleotide sequence of cDNA for sulphydryl-endopeptidase (SH-BP) from cotyledon  
A;Reference number: S05497; MUID:89386007; PMID:2780300  
A;Accession: S05497

A;Molecule type: mRNA  
A;Residues: 1-362 <AKA2>  
A;Cross-references: EMBL:X15732; NID:922061; PIDN:CAA33753.1; PID:922062  
A;Accession: S20213  
A;Molecule type: protein  
A;Residues: 132-140 <AKA3>  
R;Okamoto, T.; Nakayama, H.; Seta, K.; Isobe, T.; Minamikawa, T.  
FEBS Lett. 351, 31-34, 1994  
A;Title: Posttranslational processing of a carboxy-terminal propeptide containing a KDEL  
A;Reference number: S48684; MUID:94357273; PMID:8076688  
A;Accession: S48684  
A;Molecule type: protein  
A;Residues: 127-140/197-216;324-333;339-352 <OKA>  
C;Genetics:  
A;Introns: 150/1; 228/3; 275/3  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-362/Product: cysteine proteinase #status predicted <MAT>  
F;152,288,309/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322; DB 1; Length 362;  
Best Local Similarity 35.0%; Pred. No. 2.5e-21;  
Matches 78; Conservative 33; Mismatches 94; Indels 18; Gaps 8;

QY 8 GNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDCAS 67  
Db 126 GSVPASVDWRKGAIVTDVKDQGGCGSCWAFSTVAVEGINQIKTNKLIVLSLEQELVDCDK 185  
QY 68 Q--HGCHGDTIPRGIEYI-QHNGVQVESYRYVAREQSC--RRPNAQRFGISNYCQIYPP 122  
Db 186 EENQCGNGGLMESAFEFIKQGGITTESNPYTAQGTCDSEKVNLDLAVSIDGHENVPVN 245  
QY 123 NANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITQIRONGQPNVHNVISNAQ 181  
Db 246 DENALLKAVANQPVSAVDAAGSD---FPY-SEGVTGCDTDLN-HGVAIVGYGTVD 300  
QY 182 GVDYIWRNSDWTWNGDNGYCPAANID-----LMMIEEYP 217  
Db 301 GTNYIWRNSWGPWEQEGYIRMQRNISKEKGLCGIAMMASYP 343

## RESULT 14

S47312  
cysteine proteinase (EC 3.4.22.-) precursor - spring vetch  
C:Species: Vicia sativa (spring vetch, tare)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S51817; S47312  
R:Becker, C.; Fischer, J.; Nong, V.H.; Muenitz, K.  
Plant Mol. Biol. 26, 1207-1212, 1994  
A>Title: PCR cloning and expression analysis of cDNAs encoding cysteine proteinases from  
A:Reference number: S51817; MUID:95111103; PMID:7811978  
A:Accession: S51817  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-368 <BE2>  
A:Cross-references: UNIPROT:O41696; EMBL:X75749; NID:q535472; PIDN:CAA53377.1; PID:q5354  
A>Note: the authors translated the codon GAT for residue 241 as Glu  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:152,288,308/Active site: Cys, His, Asn #status predicted

Query Match 26.6%; Score 321; DB 2; Length 368;  
Best Local Similarity 35.1%; Pred. No. 3.1e-21;  
Matches 72; Conservative 34; Mismatches 85; Indels 14; Gaps 5;

QY 11 PAEIDLRQMTVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQELVDC--ASQ 68  
DB 129 PVHVDWRLKGAITHIKDQSGCGSCWAFSTIATVEAINKIWTGKLSLSQELVDCDRAFN 188  
QY 69 HGCHGDTIPRGTEYIQHNGVQV-ESYRYVAREQSC--RRPNAQRFGISNYCOIYPPNAN 125  
DB 189 EGCNGGLMDYAFEFIINGGGIDTDQHPYKFGEGRCDPTRKAKIVSIDGYEDVPSNNEN 248  
QY 126 KIREALAQTHSAIAVIIGIKDLDAFRH--YDGRITIQRDNGYQPNYHAVNIVGYSNAQGV 183  
DB 249 ALKKAHAQHPVSAIEASGALQLYSGVFTGCKGTSLD-----HAVVINGYSGENGL 301  
QY 184 DYWIVRNSWDTNMGDNGYGYFAANI 208  
DB 302 DYWLVRNSWGTNMGEDGYFKMERNV 326

RESULT 15  
T06416  
cysteine proteinase (EC 3.4.22.-) precursor - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T06416  
R:Lers, A.  
submitted to the EMBL Data Library, December 1997  
A:Reference number: Z15662  
A:Accession: T06416  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-466 <LER>  
A:Cross-references: UNIPROT:O49877; EMBL:AJ003137; PIDN:CAA05894.1  
A:Experimental source: Cultivar Cherry  
C:Genetics:  
A:Gene: C14  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:162,298,318/Active site: Cys, His, Asn #status predicted

Query Match 26.6%; Score 321; DB 2; Length 466;  
Best Local Similarity 36.9%; Pred. No. 4.1e-21;  
Matches 76; Conservative 28; Mismatches 82; Indels 20; Gaps 6;

QY 6 INGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQELVDC 65  
DB 134 VGDSPESIDWREKGLVGVKQDQSGCGSCWAFSAVAAMESINAIVTGNLISLSEQLVDC 193  
QY 66 --ASQHGCHGDTIPRGIEY-IQHNGVVQESYRYVAREQSC--RRPNAQRFGISNYCOIY 120  
DB 194 DRSYNEGCGGLMDYAFEFVIRKNGGIDTBEDYPYKERNVCDQYRKNKAVKVIDSYEDVP 253

QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRH-----DGRITIQRDNGYQPNYHAVNIV 175  
DB 254 VVNEKALQKAVAHQPVSIATALEAGRD---FQHYKSGIFTGCKGTAVD-----HGVVIA 303  
QY 176 GYSNAQGVVDYWIWRNSWDTNMGDNGY 201  
DB 304 GYGTENGMDYWIWRNSWGANWGENGY 329

Search completed: May 19, 2005, 17:23:28  
Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:13:36 ; Search time 60 Seconds  
(without alignments)  
1894.693 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSGNGNAPAEIDLQMR.....YFAANIDLMIEBPYVIVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1196	99.2	320	1 MMAL_DERPT	P08176 dermatophag
2	1037.5	86.0	321	1 EUMI_EURMA	P25780 euroglyphus
3	997.5	82.7	321	1 MMAL_DERFA	P16311 dermatophag
4	944.5	78.3	210	2 Q9GY70	Q9GY70 dermatophag
5	607	50.3	146	2 Q95X05	Q95X05 dermatophag
6	591.5	49.0	263	2 Q8I9P1	Q8I9P1 psoroptes o
7	469	38.9	107	2 Q95X04	Q95X04 dermatophag
8	437	36.2	133	2 Q968Y3	Q968Y3 dermatophag
9	389.5	32.3	94	2 Q7M431	Q7M431 dermatophag
10	382	31.7	221	2 Q95PJ4	Q95PJ4 blomia trop
11	341.5	28.3	333	2 Q7SX07	Q7SX07 petromyzon
12	332.5	27.6	366	2 Q6ZHP9	Q6ZHP9 oryza sativ
13	332	27.5	348	2 Q9ZQH7	Q9ZQH7 arabidopsis
14	331	27.4	326	2 Q7I7S6	Q7I7S6 callosobruc
15	330.5	27.4	393	2 Q6E7B4	Q6E7B4 brugia mala
16	328.5	27.2	326	2 Q6T857	Q6T857 fasciola gi
17	328.5	27.2	460	2 Q7XXU7	Q7XXU7 anthurium a
18	328	27.2	362	2 Q7X7A6	Q7X7A6 glycine max
19	327.5	27.2	395	1 CATL_BRUPA	O17473 brugia paha
20	327	27.1	348	1 PAP3_CARPA	P10056 carica papa
21	327	27.1	367	2 Q42673	Q42673 carica papa
22	326.5	27.1	326	2 Q24944	Q24944 fasciola he
23	326.5	27.1	327	2 Q6QXF0	Q6QXF0 agrotis seg
24	326	27.0	326	2 Q7I7S5	Q7I7S5 callosobruc
25	325	26.9	326	2 Q7I7S9	Q7I7S9 callosobruc
26	324.5	26.9	326	2 Q7JN08	Q7JN08 fasciola he
27	324	26.9	326	2 Q7I770	Q7I770 callosobruc
28	324	26.9	370	2 Q9XF79	Q9XF79 sandersonia
29	323.5	26.8	329	1 CATK_HUMAN	P43235 homo sapien
30	323.5	26.8	329	1 CATK_MACFA	P61276 macaca fasc
31	323.5	26.8	329	1 CATK_MACMU	P61277 macaca mula

32	323.5	26.8	329	2	Q6FHN2	Q6fhn2 homo sapien
33	323	26.8	323	2	Q86GF7	Q86gf7 pandalus bo
34	323	26.8	326	2	Q717S1	Q717s1 callosobruc
35	323	26.8	462	2	Q93XQ9	Q93xq9 ipomoea bat
36	322.5	26.7	326	2	Q717S4	Q717s4 callosobruc
37	322.5	26.7	326	2	Q9NGW2	Q9ngw2 fasciola gi
38	322.5	26.7	326	2	Q24137	Q24137 nicotiana t
39	322	26.7	337	2	Q7YXL3	Q7yxl3 tenebrio m
40	322	26.7	346	2	Q9LFP39	Q9lpf39 arabidopsis
41	322	26.7	362	1	CYSP_VIGMU	P12412 vigna mungo
42	321.5	26.7	445	2	Q8W182	Q8w182 brassica ol
43	321	26.6	368	2	Q41696	Q41696 vicia sativ
44	321	26.6	400	2	Q9BJM2	Q9bjm2 litomosoid
45	321	26.6	466	2	O49877	O49877 lycopersico

#### ALIGNMENTS

RESULT 1

ID	MMAL_DERPT	STANDARD;	PRT;	320 AA.
AC	P08176; Q24616;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Major mite fecal allergen Der p 1 precursor (EC 3.4.22.-) (Der p 1).			
GN	Name=DERP1;			
OS	Dermatophagoides pteronyssinus (House-dust mite).			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;			
OC	Pyroglyphidae; Dermatophagoides.			
OX	NCBI_TaxID=6956;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND POLYMORPHISM.			
RA	MEDLINE=93357682; PubMed=8353459;			
RA	Chua K.Y., Kehal P.K., Thomas W.R.;			
RT	"Sequence polymorphisms of cDNA clones encoding the mite allergen Der p I.";			
RL	Int. Arch. Allergy Immunol. 101:364-368(1993).			
RN	[2]			
RP	SEQUENCE OF 76-320 FROM N.A.			
RA	MEDLINE=88089411; PubMed=3335830;			
RA	Chua K.Y., Stewart G.A., Thomas W.R., Simpson R.J., Dilworth R.J.,			
RA	Plozza T.M., Turner K.J.;			
RT	"Sequence analysis of cDNA coding for a major house dust mite			
RT	allergen, Der p 1. Homology with cysteine proteases.";			
RL	J. Exp. Med. 167:175-182(1988).			
RN	[3]			
RP	SEQUENCE OF 81-176 FROM N.A.			
RA	MEDLINE=88114080; PubMed=3276629;			
RA	Thomas W.R., Stewart G.A., Simpson R.J., Chua K.Y., Plozza T.M.,			
RA	Dilworth R.J., Nisbet A., Turner K.J.;			
RT	"Cloning and expression of DNA coding for the major house dust mite			
RT	allergen Der p 1 in Escherichia coli.";			
RL	Int. Arch. Allergy Appl. Immunol. 85:127-129(1988).			
RN	[4]			
RP	REVISIONS TO 232-241.			
RA	MEDLINE=91215493; PubMed=2021874;			
RA	Dilworth R.J., Chua K.Y., Thomas W.R.;			
RT	"Sequence analysis of cDNA coding for a major house dust mite			
RT	allergen, Der f I.";			
RL	Clin. Exp. Allergy 21:25-32(1991).			
RN	[5]			
RP	SEQUENCE OF 99-308 FROM N.A.			
RA	MEDLINE=93130112; PubMed=1483062;			
RA	Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;			
RT	"Molecular characterization of group I allergen Eur m I from house			
RT	dust mite Euroglyphus maynei.";			
RL	Int. Arch. Allergy Immunol. 99:150-152(1992).			
RN	[6]			
RP	SEQUENCE OF 99-127.			
RA	MEDLINE=88229138; PubMed=3372999;			

RA Lind P., Hansen O.C., Horn N.;  
 RT "The binding of mouse hybridoma and human IgE antibodies to the major  
 RT T cell allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
 RT binding site location and species specificity studied by solid-phase  
 RT inhibition assays with radiolabeled antigen.";  
 RL J. Immunol. 140:4256-4262(1988).  
 RN [7]  
 RN SEQUENCE OF 99-139; 177-192; 208-224 AND 260-277, AND VARIANT ALA-222.  
 RP Pubmed=2911558;  
 RX Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;  
 RA "Structural studies on the allergen Der p 1 from the house dust mite  
 RT Dermatophagoides pteronyssinus: similarity with cysteine  
 RT proteinases.";  
 RN Protein Seq. Data Anal. 2:17-21(1989).  
 RN [8]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE=95062135; PubMed=7971950;  
 RA Topham C.M., Srinivasan N., Thorpe C.J., Overington J.P.,  
 RA Kalsheker N.A.;  
 RT "Comparative modelling of major house dust mite allergen Der p 1:  
 RT structure validation using an extended environmental amino acid  
 RT propensity table.";  
 RL Protein Eng. 7:869-894(1994).  
 CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
 CC preference for Phe or basic residues.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 CC mite allergy are bronchial asthma, allergic rhinitis and  
 CC conjunctivitis. Reacts with IgE in 80% of patients with house dust  
 CC allergy.  
 CC -!- SIMILARITY: Belongs to the peptidase C1 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U11695; AAB60215.1; -;  
 DR EMBL; M24794; AAA28296.1; ALT\_INIT.  
 DR EMBL; X65197; CAA46317.1; -;  
 DR PIR; JQ0337; JQ0337.  
 DR HSSP; P53634; IK3B.  
 DR MEROPS; C01.073; -;  
 DR InterPro; IPR000169; Pept\_cys\_actsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPAIN.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SM00845; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE HIS; 1.  
 DR Allergen; Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Polymorphism; Signal; Thiol protease; Zymogen.  
 FT SIGNAL 1 18 Potential.  
 FT PROPEP 19 98 Activation peptide.  
 FT CHAIN 99 320 Major mite fecal allergen Der p 1.  
 FT CARBOHYD 150 150 N-linked (GlcNAc...) (Potential).  
 FT ACT\_SITE 132 132 By similarity.  
 FT ACT\_SITE 268 268 By similarity.  
 FT ACT\_SITE 288 288 By similarity.  
 FT DISULFID 102 215 By similarity.  
 FT DISULFID 129 169 By similarity.  
 FT DISULFID 163 201 By similarity.  
 FT VARIANT 148 148 Y -> H.  
 FT VARIANT 179 179 E -> K.  
 FT VARIANT 222 222 V -> A.  
 FT VARIANT 234 234 S -> T.  
 FT VARIANT 313 313 E -> Q.  
 SQ SEQUENCE 320 AA; 36104 MW; A0B1P4DD09791DFE CRC64;

Query Match 99.2%; Score 1196; DB 1; Length 320;  
 Best Local Similarity 99.1%; Pred.No. 4.9e-99;  
 Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNACSGINGNAPAEIDLQMRVTPTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
 DB 99 TNACSGINGNAPAEIDLQMRVTPTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 158  
 QY 61 ELVDCASQHGCHGDTIPRGIEVIOHNGVVOESYRYVAREQSCRRRPAQRFGISNYCQIY 120  
 DB 159 ELVDCASQHGCHGDTIPRGIEVIOHNGVVOESYRYVAREQSCRRRPAQRFGISNYCQIY 218  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 180  
 DB 219 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 278  
 QY 181 QGVYDWIVRNSWDTNNGDNGYGYFAANIDLMIMEEYPYVIL 222  
 DB 279 QGVYDWIVRNSWDTNNGDNGYGYFAANIDLMIMEEYPYVIL 320  
 RESULT 2  
 EUN1 EURMA STANDARD; PRT; 321 AA.  
 ID EUN1 EURMA STANDARD; PRT; 321 AA.  
 AC P25780; O9TZ23; O9TZ24; Q9UBA0;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Mite group 1 allergen Eur m 1 precursor (EC 3.4.22.-) (Eur m I).  
 GN Name=EURM1;  
 OS Eurygllyphus maynei (Wayne's house dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Eurygllyphus.  
 OX NCBI\_TaxID=6958;  
 RN [1]  
 RP SEQUENCE FROM N.A. (EUR M 1.0101 AND EUR M 1.0102).  
 RX MEDLINE=99126275; PubMed=9925958; DOI=10.1159/000024026;  
 RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;  
 RT "Molecular analysis of the group 1 and 2 allergens from the house dust  
 RT mite, Eurygllyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 118:15-22(1999).  
 RN [2]  
 RP SEQUENCE OF 99-309 FROM N.A.  
 RX MEDLINE=93130112; PubMed=1483062;  
 RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
 RT "Molecular characterisation of group 1 allergen Eur m I from house  
 RT dust mite Eurygllyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 99:150-152(1992).  
 CC -!- FUNCTION: Probable thiol protease.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 CC mite allergy are bronchial asthma, allergic rhinitis and  
 CC conjunctivitis.  
 CC -!- SIMILARITY: Belongs to the peptidase C1 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF047610; AAC82351.1; -;  
 DR EMBL; AF047611; AAC82352.1; ALT\_INIT.  
 DR EMBL; AF047612; AAC82353.1; -;  
 DR EMBL; X60073; CAA42677.1; -;  
 DR PIR; S21864; S21864.  
 DR HSSP; P53634; IK3B.  
 DR MEROPS; C01.073; -;  
 DR InterPro; IPR000169; Pept\_cys\_actsite.

```

DR InterPro; IPR000668; Peptidase Cl.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PF00705; PAPAIN.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept Cl; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOI_PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE HIS; 1.
KW Allergen; Glycoprotein; Hydrolase; Signal; Thiol protease; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 98
FT CHAIN 99 321 Mite group 1 allergen Eur m 1.
FT ACT_SITE 133 133 By similarity.
FT ACT_SITE 269 269 By similarity.
FT ACT_SITE 289 289 By similarity.
FT DISULFID 130 170 By similarity.
FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT VARIANT 36 36 T -> S (in Eur m 1.0102).
FT VARIANT 126 126 M -> N (in Eur m 1.0102).
FT VARIANT 320 320 M -> I (in Eur m 1.0102).
SQ SEQUENCE 321 AA; 36290 MW; 6CFD44FEC725999E CRC64;

Query Match 86.08; Score 1037.5; DB 1; Length 321;
Best Local Similarity 83.94; Pred. No. 8.4e-85;
Matches 187; Conservative 19; Mismatches 16; Indels 1; Gaps 1;

QY 1 TNACSG-NAPAEIDLQRMTVPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAE 59
DB 99 TYACINSVSLPSLDLSRLTVPPIRMQCGCGSCWAFSGVASTESAYLAHRNOSLDLAE 158

QY 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFISNYCQI 119
DB 159 QELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRYGLKNYCQI 218

QY 120 YPPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 179
DB 219 SPDSNKRQALQTHTAVAVIIGIKDLNAPRHYDGRITIMQHDNGYQPNYHVNIVGYSN 278

QY 180 AQGVYDVIWRNSWDTNMGDNGYGYFAANIDLMIEEYPPYVVL 222
DB 279 TQGVYDVIWRNSWDTNMGDNGYGYFAANINLMIEQYPPYVVL 321

RESULT 3
MMAL DERPA
ID _MMAL_DERPA STANDARD; PRT; 321 AA.
AC P16311;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Major mite fecal allergen Der f 1 precursor (EC 3.4.22.-) (Der f I).
GN Name=DERf1;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmatida; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91215493; PubMed=2021874;
RA Dilworth R.J.; Chua K.Y.; Thomas W.R.;
RT "Sequence analysis of cDNA coding for a major house dust mite
RT allergen, Der f I."
RL Clin. Exp. Allergy 21:25-32 (1991).
RN [2]
RP SEQUENCE OF 98-309 FROM N.A.
RA Kent N.; Hill M.R.; Keen J.N.; Holland P.W.; Hart B.J.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 99-128.
RX MEDLINE=88229138; PubMed=3372999;
RA Lind P.; Hansen O.C.; Horn N.;

```

```

RT "The binding of mouse hybridoma and human IgE antibodies to the major
RT fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative
RT binding site location and species specificity studied by solid-phase
RT inhibition assays with radiolabeled antigen.";
RL J. Immunol. 140:4256-4262(1988).
CC -I- FUNCTION: Thiol protease that hydrolyzes proteins, with a
CC preference for Phe or basic residues.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- ALLERGEN: Causes an allergic reaction in human. Common symptoms of
CC mite allergy are bronchial asthma, allergic rhinitis and
CC conjunctivitis.
CC -I- SIMILARITY: Belongs to the peptidase C1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X65196; CAA46316.1; -.
CC PIR; A27634; A27634.
CC HSP; P53634; IK3B.
CC MEROPS; C01.073; -.
CC InterPro; IPR000169; Pept_cys_actsite.
CC InterPro; IPR000668; Peptidase_C1.
CC Pfam; PF00112; Peptidase_C1; 1.
CC PRINTS; PR00705; PAPAIN.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SM00645; Pept Cl; 1.
CC PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
CC PROSITE; PS00139; THIOI_PROTEASE CYS; 1.
CC PROSITE; PS00639; THIOI_PROTEASE HIS; 1.
KW Allergen; Direct protein sequencing; Glycoprotein; Hydrolase; Signal;
KW Thiol protease; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 98 Activation peptide.
FT CHAIN 99 321 Major mite fecal allergen Der f 1.
FT ACT_SITE 133 133 By similarity.
FT ACT_SITE 269 269 By similarity.
FT ACT_SITE 288 288 By similarity.
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT DISULFID 130 170 By similarity.
FT DISULFID 164 202 By similarity.
FT CONFLICT 201 201 R -> Q (in Ref. 2).
FT CONFLICT 282 282 D -> V (in Ref. 2).
SQ SEQUENCE 321 AA; 36435 MW; 04523E54EBB476E CRC64;

Query Match 82.74; Score 997.5; DB 1; Length 321;
Best Local Similarity 81.24; Pred. No. 3.3e-81;
Matches 181; Conservative 18; Mismatches 23; Indels 1; Gaps 1;

QY 1 TNACSG-NAPAEIDLQRMTVPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAE 59
DB 99 TSACRINSVNPSELDLSRLTVPPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAE 158

QY 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFISNYCQI 119
DB 159 QELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRYGLKNYCQI 218

QY 120 YPPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 179
DB 219 YPPDVKQIREALQTHTAVAVIIGIKDLNAPRHYDGRITIMQHDNGYQPNYHVNIVGYSN 278

QY 180 AQGVYDVIWRNSWDTNMGDNGYGYFAANIDLMIEEYPPYVVL 222
DB 279 TQGVYDVIWRNSWDTNMGDNGYGYFAANINLMIEQYPPYVVL 321

RESULT 4
Q9GY0

```

```

ID Q9GY0 PRELIMINARY; PRT; 210 AA.
AC Q9GY0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Allergen Der fi (Fragment).
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Hao M.Q., Xu J., Zhong N.S.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family C1.
DR EMBL; A285763; AAG00520.1; -.
DR PIR; A27634; A27634.
DR HSSP; P80067; 1JQP.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE HIS; UNKNOWN_1.
KW Hydrolase; Protease; Thiol protease.
FT NON TER 1
FT NON TER 210
SQ SEQUENCE 210 AA; 23548 MW; BA08029D642EB990 CRC64;

Query Match 78.3%; Score 944.5; DB 2; Length 210;
Best Local Similarity 81.4%; Pred. No. 1.1e-76;
Matches 171; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY 2 NACSING-NAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 SACRINSVNPSELDRLSUKTVPTRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLSEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNAQRFSGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQQNGVVEERSYPVAREQQCRRENSQHYGISNYCQIY 120

QY 121 PPNAKIREALATHSAIAVIIGIKOLDAPRHVDGRTTIQRDNGYQPNYHVNIVGYSNA 180
DB 121 PPDVKQIREALTQHTAIAVIIGIKOLRAFQHYDGRITIIQHDNGYQPNYHVNIVGYGST 180

QY 181 QGVYDIWVRNSWDTNNGDNGYGFANIDL 210
DB 181 QGVYDIWVRNSWDTTWGDSGYGFQAGNNL 210

RESULT 5
Q95X05 PRELIMINARY; PRT; 146 AA.
AC Q95X05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cysteine proteinase (Fragment).
GN Names=CPW2;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF194431; AAL14424.1; -.
DR HSSP; P43235; 1BY8.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
FT NON TER 1
FT NON TER 146
SQ SEQUENCE 146 AA; 16852 MW; BB304800946D4047 CRC64;

Query Match 50.3%; Score 607; DB 2; Length 146;
Best Local Similarity 84.2%; Pred. No. 1.5e-46;
Matches 112; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 39 GVAATESAYLAHRNOSLDLAEQELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVA 98
DB 1 GVAATESAYLAHRNOSLDLSEQELVDCASQHGCHGDTIPRGIEYIQQNGVVEERSYPVA 60

QY 99 RQSCRRPNAQRFSGISNYCQIYPPNANKIREALATHSAIAVIIGIKOLDAPRHVDGRTTI 158
DB 61 RQSCRRPNSQHYGISNYCQIYPPDVVKQIREALTQHTAIAVIIGIKOLDAPRHVDGRTTI 120

QY 159 IORDNGYQPNYHA 171
DB 121 IQHDNGYQPNYHA 133

RESULT 6
Q819P1 PRELIMINARY; PRT; 263 AA.
AC Q819P1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Derp1 antigen (Fragment).
OS Psoroptes ovis (Sheep scab mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Sarcotoides;
OC Psoroptidae; Psoroptes.
OX NCBI_TaxID=83912;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22294998; PubMed=12406195;
RA Lee A.J., Machell J., Van Den Broek A.H.M., Nisbet A.J.,
RA Miller H.R.P., Isaac R.E., Huntley J.F.;
RT "Identification of an antigen from the sheep scab mite, Psoroptes
RT ovis, homologous with house dust mite group 1 allergens.";
RL Parasite Immunol. 24:413-422(2002).
DR EMBL; AF495854; AAO14671.1; -.
DR HSSP; P14080; 1YAL.
DR MEROPS; C01.073; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
FT NON TER 1
FT NON TER 263
SQ SEQUENCE 263 AA; 29576 MW; BF6DD21006DAB5B0 CRC64;

Query Match 49.0%; Score 591.5; DB 2; Length 263;
Best Local Similarity 64.4%; Pred. No. 7.2e-45;
Matches 112; Conservative 22; Mismatches 39; Indels 1; Gaps 1;

QY 3 ACSING-NAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 61
DB 90 ACOIGAVNIPNEIDLRLGVTIKIKQVACGSCWAFSGVATVESNVLSDNLSEQ 149

QY 62 LVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNAQRFSGISNYCQIYP 121
|||||
```

```
Db 150 LVDCASQHGCGDVTVLNGLRYIQKNGVVEQSYPYKAREGRCQPNKRYGIKDLQIYP 209
Qy 122 PNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIV 175
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 210 PNGDKIRTYLATKQAALSIIIGIRDLDSFRHYDGRITIIQSDNGGKGRNFHAINIV 263

RESULT 7
Q95X04 PRELIMINARY; PRT; 107 AA.
AC Q95X04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cysteine proteinase (Fragment).
GN Name=CPW3;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Kim K.Y., Park S.K., Yun H.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194432; AAL14425.1; -.
DR InterPro; IPR00169; Pept_cys_acsite.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWNW_1.
FT NON_TER 1
FT NON_TER 107
FT NON_TER 107
SQ SEQUENCE 107 AA; 12277 MW; A80E7876CBA6F97A CRC64;

Query Match 38.9%; Score 469; DB 2; Length 107;
Best Local Similarity 80.4%; Pred. No. 2.5e-34;
Matches 86; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 77 PRGIEYQHNGVQESYRYVAREQSCRPNQAQFGISNYCQIYPPNANKIREALAOHTS 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 PRGIEYQHNGVVEERSYFYVAREQCCRRPSQSHYGISNYCQIYPPDPVKQIREALTQHT 60

Qy 137 AIAVLIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNAQGV 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AIAVLIGIKDLRAFOHYDGRITIIQRDNGYQPNYHAVNIVGYSTQGV 107

RESULT 8
Q968Y3 PRELIMINARY; PRT; 133 AA.
AC Q968Y3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cysteine proteinase (Fragment).
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Yun H.C., Kim K.Y., Park S.Y., Park S.K.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF145247; AAK38773.1; -.
DR HSSP; P60994; 1IWD.
DR MEROPS; C01.073; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
FT NON_TER 1
FT NON_TER 133
FT NON_TER 133
SQ SEQUENCE 133 AA; 14965 MW; 5033C26B15E68E9C CRC64;
```

```
Query Match 36.2%; Score 437; DB 2; Length 133;
Best Local Similarity 61.4%; Pred. No. 2.4e-31;
Matches 81; Conservative 17; Mismatches 34; Indels 0; Gaps 0;

Qy 31 CGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDCASQHGCHGDTIPRGIEYQHNGVQ 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CGSCWAFAGVAATESAYLAHRNQSLLDAEQELVDCARRGCHGDTIPRGLYIQNGIVE 60

Qy 91 ESYRYVAREQSCRPNQAQFGISNYCQIYPPNANKIREALAOHTSAIAVIIGIKDLDAF 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 EQAYEYNARENCCPPENPRHSIQYQCIDHSNVELIKTALDKYKSAVAVIINHINAF 120

Qy 151 RHYDGRITIIQRD 162
|||||:|||||:
Db 121 RHYDGSYVITTD 132

RESULT 9
Q7M431 PRELIMINARY; PRT; 94 AA.
AC Q7M431;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major fecal allergen Der p I (Fragments).
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE.
RX MEDLINE=89098855; PubMed=2911558;
RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;
RT "Structural studies on the allergen Der p I from the house dust mite
RT Dermatophagoides pteronyssinus: similarity with cysteine
RT proteinases.";
RL Protein Seq. Data Anal. 2:17-21 (1989).
DR PIR; S03380; S03380.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR ProDom; PD000158; Peptidase_C1; 1.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10327 MW; 9BF744165C8428A8 CRC64;

Query Match 32.3%; Score 389.5; DB 2; Length 94;
Best Local Similarity 48.6%; Pred. No. 2.9e-27;
Matches 88; Conservative 0; Mismatches 4; Indels 89; Gaps 4;

Qy 1 TNACISINGNAPAEIDLQRMRTVTPIR--MOGCGSGCWAFSGVAATESAYLAHRNQSLLDA 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 TNACISINGNAPAEIDLQRMRTVTPIR--MOGCGSGCWAFSGVAATESAYLAHRNQSLLDA 58

Qy 59 EQELVDCASQHGCHGDTIPRGIEYQHNGVQESYRYVAREQSCRPNQAQFGISNYCQ 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 44 -----GIEYQHNGVQESY-----RFGISNYCQ 68

Qy 119 IYPPNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYS 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 IYPPNANK-----DNGYQPNYXAVNIVGYX 93

Qy 179 N 179
Db 94 N 94

RESULT 10
Q95PJ4 PRELIMINARY; PRT; 221 AA.
AC Q95PJ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```





Best Local Similarity 35.7%; Pred. No. 1.9e-21;  
Matches 79; Conservative 34; Mismatches 91; Indels 17; Gaps 7;

Qy 9 NAPAEDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDCAS- 67  
Db 142 NLPWAVDRKKGATVPKNQGGCGSCWAFSGVAATEVGINQIVTKVLSLSEQELMDCDNT 201  
Qy 68 -QHCHGDTIPRGIEYIQHNVGVQESYRYRYVAREQSCR--RPNARFGISNYCQIYPPN 123  
Db 202 FNHCGRGLMDFAFAYIMGNQGIYTEDYPYLMEEGYCREKQPHSKVITITGYEDVPANS 261  
Qy 124 ANKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQYNYHVNIVGYSNAQGV 183  
Db 262 ETSLLKALAHQPVSVGTAAGSRD---FQYKYG-GIFDGEIGIQPD-HALTAVGYGSIYQ 316  
Qy 184 DYTWIRNSWDTNMGNGY-----GYFAANIDLMIEEYP 217  
Db 317 DYIIMKNSWGNWGEQGYFRIRRTGKPEGVCDIYKIASYP 357

RESULT 13

Q9ZQH7  
ID Q9ZQH7 PRELIMINARY; PRT; 348 AA.  
AC Q9ZQH7  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Cysteine proteinase.  
GN Name=At2g27420;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,  
RA Barnstead M.E., Mason T.M., Bowman C.L., Renning C.M., Benito M.-I.,  
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,  
RA Fraser C.M., Venter J.C.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Town C.D., Kaul S.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
RA Davis R.W., Ecker J.R., Theologis A.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; AC006232; AAD15594.1; -  
DR EMBL; AY064033; AAL36389.1; -  
DR EMBL; AY096388; AAM20029.1; -  
DR F84672; F84672.  
DR HSP; P53634; IK3B.  
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
DR PROSITE; PS00639; THIOI\_PROTEASE HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Thiol protease.  
SQ SEQUENCE 348 AA; 38738 MW; EB86ABEC2B553E76 CRC64;

Query Match 27.5%; Score 332; DB 2; Length 348;  
Best Local Similarity 36.2%; Pred. No. 1.9e-21;  
Matches 84; Conservative 32; Mismatches 86; Indels 30; Gaps 11;

Qy 5 SINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVD 64  
Db 126 SDNGES---MDWRQEGAVTPVKYQRCGCGWAFSAVAAREGITTKITKGELVSLSEQQLLD 182  
Qy 65 CASQH--GCHGDTIPRGIEY-IOHNGVQESYRYRYVAREQSCRPN------QRFGISNY 116  
Db 183 CDRDYNQCGRGIMSKAFVILIKNGITTEDNTPYQESQTCSSSTLSSSFRAATISGY 242  
Qy 117 CQYPPNANKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGY--QPNYHAVN 173  
Db 243 -ETVPMNN---EEALLQAVSQPVSVGIEGTGAFRHYSQGVF---NGECGTDLHHAVT 294  
Qy 174 IVGYS-NAQGVYVIVNSNDTWGNGYGYFAANID------LMIEEYP 217  
Db 295 IVGYMSEEGTKYVWVKNWNGETWNGYWRKRDVDAPOGMCGLAILAFYP 346

RESULT 14

Q717S6  
ID Q717S6 PRELIMINARY; PRT; 326 AA.  
AC Q717S6;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Putative gut cathepsin L-like cysteine protease.  
GN Name=CPA15;  
OS Callosobruchus maculatus (cowpea weevil).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
OC Phytophaga; Chrysomeloidea; Chrysomelidae; Bruchinae; Bruchini;  
OC Callosobruchus.  
OX NCBI\_TaxID=64391;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhu-Salzman K., Ahn J.-E., Koiba H., Salzman R.A.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; AF544838; AAQ11969.1; -  
DR HSP; P43235; IAYU.  
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
DR PROSITE; PS00639; THIOI\_PROTEASE HIS; 1.  
KW Hydrolase; Protease; Thiol protease.  
SQ SEQUENCE 326 AA; 36653 MW; A9265E018FFD2A8D CRC64;

Query Match 27.4%; Score 331; DB 2; Length 326;  
Best Local Similarity 32.7%; Pred. No. 2.2e-21;  
Matches 71; Conservative 38; Mismatches 94; Indels 14; Gaps 6;

```
QY 14 IDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQLVDCASQ----H 69
Db 116 VDMREGAVTPVKDQANCGSCWAFSAVGATEGQFFKNGTFLVLSAQELVDCATEDYGN 175
QY 70 GCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRPNQORFGISNYCQIYPNPANKIRE 129
Db 176 GCKGGLMGQAFDFVQDEGIQTEESYPVEGRSSCKSGEYVTKVTKY--VFPLDEQEMAR 233
QY 130 ALAQTHSAIAVIIGIKDLDAFRHYDRTTIQR---DNGYQPNYHVNIVGYSNAQGVYDW 186
Db 234 TVA-AKGPVAVATEASQLS---FYDKGIVDRCRCNKREDLNHGLVVLVVGSGSENGVDY 289
QY 187 IVNSWDTNWGDNGYGYFAANIDLMLIEEY-PYVVIL 222
Db 290 IVKNSWGDWGEKGYFLKKDKVAKCGIGYNYNPYILL 326
```

## RESULT 15

```
Q6E7B4 PRELIMINARY; PRT; 393 AA.
AC Q6E7B4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Cathepsin L-like cysteine proteinase.
GN Name=CPL-4;
OS Brugia malayi (Filarial nematode worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A.
RA Guillian D.B., Hong X., McKerrow J.H., Blaxter M.L., Oksov Y., Liu J.,
RA Ghedin E., Lustigman S.;
RT "A gene family of cathepsin L-like proteases of filarial nematodes are
RT associated with larval molting and cuticle and eggshell remodeling.";
RL Mol. Biochem. Parasitol. 136:227-242(2004).
CC -|- SIMILARITY: Belongs to peptidase family C1.
DR EMBL; AY533166; AAU07057.1; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase C1.
DR InterPro; IPR000169; Pept. cys acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00640; THIOLEPROTEASE ASN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE CYS; 1.
KW Hydrolase; Protease; Thiol protease.
SQ SEQUENCE 393 AA; 44527 MW; C0A34546595B1E93 CRC64;
```

```
Query Match 27.4%; Score 330.5; DB 2; Length 393;
Best Local Similarity 37.9%; Pred. No. 3.1e-21;
Matches 77; Conservative 33; Mismatches 83; Indels 11; Gaps 7;
QY 11 PAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQLVDCASQ---S 67
Db 181 PDLVDMRTKGAIVFVRNQSGCGYAFATAALEAYHKQMTGRLLDLSPQNIIVDCTWDLG 240
QY 68 QHCGHGTIPRGIEYIQHNGVQVESYRYVAREQSCR-RPNAQORFGISNYCQIYPNPANK 126
Db 241 NKCGNGYMTAFQYASRYGIAMESYPYVGTQRQRWQNIIVTNDGNEIQPQDELA 300
QY 127 IREALAQTSAIAVIIGIKDLDAFRHY-DGRTTIQRDNGYQPNYHVNIVGYSNAQGV-D 184
Db 301 LKHAVAKRGPVVVGISGSK--RSFRFYKDG--VYSEGNCGRPD-HAVLVVVGYGTDRSYGD 355
QY 185 YWIVRNSWDTNWGDNGYGYFAAN 207
Db 356 YWIVRNSWDTNWGDNGYGYFAAN 378
```

Search completed: May 19, 2005, 17:22:44  
Job time : 62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:12:26 ; Search time 164 Seconds  
(without alignments)  
523.541 Million cell updates/sec

Title: US-09-867-159a-2  
Perfect score: 1206  
Sequence: 1 TNACSSINGNAPAEIDLQRM.....YFAANIDLMIEPPYVIL 222

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206	100.0	222	5	AAO20568 Cysteine
2	1206	100.0	222	5	ABB98533 Cysteine
3	1206	100.0	245	2	AAR22433 Der p I a
4	1206	100.0	245	2	AAR39359 Dermatoph
5	1206	100.0	245	2	AAR51727 Der p I.
6	1206	100.0	245	2	AAR47063 Protein a
7	1206	100.0	245	2	AAR66545 Der p I a
8	1206	100.0	245	2	AAW71908 Dermatoph
9	1206	100.0	245	2	AAW50356 Dermatoph
10	1206	100.0	245	4	AAU18959 House dus
11	1206	100.0	245	5	ABG71812 European
12	1206	100.0	320	2	AAR49920 Protein a
13	1203	99.8	222	4	AAU07748 House dus
14	1200	99.5	222	4	AAB98347 D. pteron
15	1200	99.5	302	4	AAB98346 D. pteron
16	1200	99.5	302	5	AAW50623 Recombina
17	1200	99.5	320	8	ADK52158 Full leng
18	1199	99.4	362	8	ADM57302 Modular a
19	1196	99.2	222	4	AAU07746 House dus
20	1196	99.2	222	5	ABG67023 House dus
21	1196	99.2	222	8	ADK52140 Der p I a
22	1196	99.2	222	8	ADR87225 Dust mite
23	1196	99.2	223	4	AAB98345 D. pteron
24	1196	99.2	223	4	AAW98331 D. pteron
25	1196	99.2	282	8	ADL27398 Polypepti

26	1196	99.2	302	4	AAB98348 D. pteron
27	1196	99.2	302	4	AAB98330 D. pteron
28	1196	99.2	302	6	AAE36747 Dermatoph
29	1196	99.2	302	8	ADR97692 European
30	1196	99.2	303	4	AAB98344 D. pteron
31	1196	99.2	320	2	AAV25580 D. pteron
32	1196	99.2	320	2	AAV08592 D. pteron
33	1196	99.2	320	4	AAB98329 D. pteron
34	1196	99.2	320	6	ABB80128 Der p I. 6
35	1196	99.2	320	6	ABP98482 Amino aci
36	1196	99.2	320	7	ADC34830 House dus
37	1196	99.2	320	7	ADE38098 European
38	1196	99.2	320	8	ADM57314 Modular a
39	1196	99.2	320	8	ADQ14389 European
40	1196	99.2	320	8	ADS52096 Major mite
41	1196	99.2	320	8	ADS14367 Dust mite
42	1195	99.1	339	2	AAV08596 D. pteron
43	1191	98.8	320	6	ABU11147 House dus
44	1191	98.8	343	2	AAV08597 D. pteron
45	1189	98.6	222	4	AAU07747 House dus

ALIGNMENTS

RESULT 1  
AAO20568  
ID AAO20568 standard; protein; 222 AA.  
XX  
AC AAO20568;  
XX  
DT 02-JAN-2003 (first entry)  
XX  
DE Cysteine protease protein.  
XX  
KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 105  
FT /note= "Encoded by ACC"  
XX  
PN WO200278736-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-FR001098.  
XX  
PR 30-MAR-2001; 2001FR-00004370.  
PR 03-MAY-2001; 2001FR-00005929.  
PR 29-MAY-2001; 2001US-00867159.  
XX  
PA (ANTI-) ANTIALIS S.A.R.L.  
XX  
PI Loria E, Terrasse G, Trehin Y;  
XX  
DR WPI; 2002-750636/81.  
DR N-PSDB; AAL41281.  
XX  
PT Antiallergic compositions containing an anti-histamine, a histamine  
PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
PT for the allergen.  
XX  
PS Claim 13; Page 30-31; 32pp; French.  
XX  
CC The invention relates to antiallergic compositions containing an anti-  
CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
CC isolated nucleic acid molecule that has at least one polynucleotide  
CC sequence coding for the allergen, together with a pharmaceutical carrier.

CC The pharmaceutical composition of the invention is useful as a non-specific anti-allergic treatment, and also useful in the treatment of allergic hypersensitivity, allergic asthma, allergic rhinitis, and allergic atopic eczema. This sequence represents the cysteine protease protein relating to the anti-allergic compositions of the invention

XX  
XX  
SQ Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.7e-127;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSLNGAPAEIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
DB 1 TNACSLNGAPAEIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120

QY 121 PPNAKIREALQTHSAIAVIIGIKOLDAPRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
DB 121 PPNAKIREALQTHSAIAVIIGIKOLDAPRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222  
DB 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222

RESULT 2  
ABB98533  
ID ABB98533 standard; protein; 222 AA.  
AC ABB98533;  
XX  
XX  
DT 13-DEC-2002 (first entry)  
DE Cysteine protease.  
XX  
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological; immunotherapy; allergen; allergic hypersensitivity reaction;  
KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
KW cysteine protease.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 105 /note= "Encoded by CCA"  
FT  
FT  
FT  
PN FR2822709-A1.  
XX  
XX  
PD 04-OCT-2002.  
XX  
PF 03-MAY-2001; 2001FR-00005929.  
XX  
PR 30-MAR-2001; 2001FR-00004370.  
XX  
PA (ANTI-) ANTIALIS SARL.  
XX  
XX Loria E, Terrasse G, Trehin Y;  
XX WPI; 2002-735037/80.  
DR N-PSDB; ABQ80833.  
XX  
XX Antiallergic composition, useful for preventing and treating e.g. asthma, rhinitis or eczema, containing at least two of allergen, antihistamine and histamine synthesis inhibitor.  
XX  
XX Claim 7; Page 27-28; 33pp; French.  
XX  
XX The present invention relates to an anti-allergic pharmaceutical

CC composition (I) comprising a pharmaceutical carrier containing an active agent combination of at least two of: an allergen; an antihistamine; and a histamine synthesis inhibitor. (I) is used for treating or preventing allergic hypersensitivity reactions, especially allergic asthma, allergic rhinitis or allergic atopic eczema, in babies, children or adults. The present sequence is cysteine protease from Dermatophagoides pteronyssinus, which was used as an allergen in the invention

XX  
SQ Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.7e-127;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSLNGAPAEIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
DB 1 TNACSLNGAPAEIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120

QY 121 PPNAKIREALQTHSAIAVIIGIKOLDAPRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
DB 121 PPNAKIREALQTHSAIAVIIGIKOLDAPRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222  
DB 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222

RESULT 3  
AAR22433  
ID AAR22433 standard; protein; 245 AA.  
XX  
XX  
AC AAR22433;  
XX  
DT 27-AUG-2003 (revised)  
DT 04-AUG-1992 (first entry)  
XX  
DE Der p I allergen.  
XX  
KW House dust mite; desensitisation therapy; sensitivity.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
FH Key Location/Qualifiers  
FT Cleavage-site 7..8 /note= "signal peptide cleavage"  
FT Modified-site 68..70 /note= "N-glycosylation site"  
FT  
XX  
PN WO9204445-A.  
XX  
XX  
PD 19-MAR-1992.  
XX  
PF 10-SEP-1991; 91WO-AU000417.  
XX  
PR 11-SEP-1990; 90US-00580655.  
XX  
PA (WAUR-) WEST AUST RES INST.  
XX  
XX Thomas WR, Chua KY;  
PI  
XX WPI; 1992-114353/14.  
DR N-PSDB; AAQ23373.  
XX  
XX DNA encoding allergens of house dust mite Dermatophagoides - and isolated protein allergens and peptide(s) useful in treating and detecting sensitivity to mites, esp. D. farinae.  
XX  
XX Claim 14; Fig 7; 53pp; English.  
PS  
XX

CC The protein sequence was deduced from the cDNA sequence obtd. by  
 CC screening a dermatophagoides cDNA library with two probes comprising the  
 CC Der p 1 cDNA BamHI fragments 1-348 and 349-857. The protein allergen is  
 CC used as a diagnostic reagent in detecting and treating sensitivity to  
 CC house dust mites and in desensitisation therapy. See also AAR22431,2.  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 60  
 DB 24 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNAPQFGISNYCOIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNAPQFGISNYCOIY 143  
 QY 121 PPNAKIREALAOATHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHNAVNIYGSNA 180  
 DB 144 PPNAKIREALAOATHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHNAVNIYGSNA 203  
 QY 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222  
 DB 204 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 4  
 AAR39359  
 ID AAR39359 standard; protein; 245 AA.  
 AC AAR39359;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 12-AUG-1993 (first entry)  
 XX  
 XX Dermatophagoides pteronyssinus Dep p1.  
 XX  
 XX T cell epitope; house dust mite; allergy; soluble; allergen.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 XX  
 XX WO9308279-A1.  
 XX  
 XX 29-APR-1993.  
 XX  
 XX 15-OCT-1992; 92WO-US008637.  
 XX  
 XX 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 PI  
 XX WPI; 1993-152472/18.  
 DR N-PSDB; AAQ41211.  
 DR  
 XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 PT  
 XX Disclosure; Page 47; 176pp; English.  
 PS  
 XX The cDNA encoding Dermatophagoides pteronyssinus group I allergen Dep p1  
 CC was obtd. in plasmid form as subclone from lambda gt11 (Chua et al., J.  
 CC Exp. Med. 167: 175-182 (1988)). The cDNA was used to express a Dep p1  
 CC allergen which was capable of raising an immune response when contacted  
 CC with anti-Dep p1 antibodies. The protein was used to design a series of  
 CC overlapping peptides synthesised by standard techniques to cover the  
 CC whole Dermatophagoides pteronyssinus Dep p1 sequence. The T cell epitopes

CC of the protein were mapped by detection of the peptide's ability to  
 CC stimulate T cell activity. The peptides may be used for diagnosis and  
 CC treatment of sensitivity to house dust mite allergens. When administered  
 CC to house dust mite sensitive individuals, the peptides are capable of  
 CC modifying the allergic response to the allergens. The peptides may be  
 CC modified for e.g. increasing solubility, enhancing therapeutic or  
 CC preventive efficacy or stability. See also AAR34686-700, AAR36398-490 and  
 CC AAR39360-2. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 60  
 DB 24 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNAPQFGISNYCOIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNAPQFGISNYCOIY 143  
 QY 121 PPNAKIREALAOATHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHNAVNIYGSNA 180  
 DB 144 PPNAKIREALAOATHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHNAVNIYGSNA 203  
 QY 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222  
 DB 204 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 5  
 AAR51727  
 ID AAR51727 standard; protein; 245 AA.  
 AC AAR51727;  
 XX  
 DT 01-FEB-1995 (first entry)  
 XX  
 XX Der p I.  
 XX  
 XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerance.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 OS  
 XX ZA9302677-A.  
 PN  
 XX 26-JAN-1994.  
 PD  
 XX 16-APR-1993; 93ZA-00002677.  
 PF  
 XX 16-APR-1993; 93ZA-00002677.  
 PR  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 PI  
 XX WPI; 1994-126807/15.  
 DR N-PSDB; AAQ62308.  
 DR  
 XX Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 PT  
 XX Disclosure; Page 48-49; 154pp; English.  
 PS  
 XX

CC This sequence represents the group I protein allergen from the house dust  
 CC mite *D. pteronyssinus*, Der p I. The Der p I protein shows high homology  
 CC to the group I protein allergens derived from *D. farinae*, Der f I, having  
 CC an identity of 81%. Fragments of these proteins, and the corresponding  
 CC group II allergens, Der p II and Der f II, (see AAR51731-841) represent T  
 CC - cell epitopes. Fusion peptides may be produced which comprise at least  
 CC two or these antigenic fragments. Each region of these peptides may be  
 CC derived from the same, or different, mite allergens. The antigenic  
 CC fragments may be altered by substitution, deletion or addition to enhance  
 CC their antigenicity. These peptides may be produced by chemical synthesis,  
 CC chemical cleavage of the protein allergen or by recombinant techniques.  
 CC These peptides, when administered to a house dust mite sensitive  
 CC individual, are capable of modifying the allergic response of the  
 CC individual to the allergen. These peptides do not bind to immuno- globulin  
 CC E (IgE), or bind IgE to a lesser extent than the full length protein  
 CC allergen. This reduces the major complications of standard immunotherapy,  
 CC which are IgE-mediated responses such as anaphylaxis. Exposure of mite  
 CC allergic patients to these peptides may tolerate or anergise appropriate  
 CC T cell subpopulations such that they become unresponsive to mite  
 CC allergens and do not participate in mounting an immune response upon  
 CC exposure. Administration of the peptide may also modify the lymphokine  
 CC secretion profile as compared with exposure to the naturally occurring  
 CC mite protein allergen  
 CC  
 CC Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 24 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 143  
 QY 121 PPNAKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 144 PPNAKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
 QY 181 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
 DB 204 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 245

RESULT 6  
 AAR47063  
 ID AAR47063 standard; protein; 245 AA.

XX AC AAR47063;  
 XX  
 XX 25-MAR-2003 (revised)  
 XX 17-OCT-1994 (first entry)  
 XX Protein allergen of Derp I.  
 XX  
 XX Der pI; House Dust Mite Allergen.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 XX  
 XX Key Location/Qualifiers  
 XX Protein 1..222 /label= Der pI preproenzyme  
 XX Cleavage-site 9..111 /label= Cleavage site  
 XX /note= "proenzyme remains"  
 XX Cleavage-site 11..23 /label= cleavage site  
 XX /note= "cleavage between pro-Der pI and pre-Der "

XX WO9405790-A1.  
 PN

XX 17-MAR-1994.  
 XX  
 XX 10-SEP-1993; 93WO-US008518.  
 XX  
 XX 10-SEP-1992; 92US-00945288.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 XX Thomas WR, Chua K;  
 XX WPI; 1994-101195/12.  
 XX N-PSDB; AAO58665.  
 XX  
 XX New protein allergens of house dust mite - used for diagnosing and  
 XX treating sensitivity in an individual to house dust mite allergens.  
 XX  
 XX Example 1; Fig 1; 98pp; English.  
 XX  
 XX AAR47063 is a preproenzyme Der pI. The amino acid sequence preceding the  
 XX mature protein sequence contains cleavage sites for the pre- and  
 XX proenzyme forms, with residues 1-11 corresponding to a partial signal  
 XX peptide sequence. The mature protein can be used to detect sensitivity in  
 XX an individual to house dust mite and to reduce the sensitivity of the  
 XX individual. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 24 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAREQSCRRPNAQRFGISNYCQIY 143  
 QY 121 PPNAKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 144 PPNAKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
 QY 181 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
 DB 204 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 245

RESULT 7  
 AAR66545  
 ID AAR66545 standard; protein; 245 AA.

XX AC AAR66545;  
 XX  
 XX 25-MAR-2003 (revised)  
 XX 31-JUL-1995 (first entry)  
 XX Der p I allergen.  
 XX  
 XX Der p I; allergen; antigen; immunological tolerance; peptide;  
 XX KW cryptic epitope.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 XX  
 XX WO9427634-A1.  
 XX  
 XX 08-DEC-1994.  
 XX  
 XX 01-JUN-1994; 94WO-AU000292.  
 XX  
 XX 02-JUN-1993; 93US-00072832.  
 XX  
 XX

```

PA (CHIL-) INST CHILD HEALTH RES.
XX Thomas WR;
XX WPI; 1995-022467/03.
XX N-PSDB; AAQ79618.
XX Inducing immunological tolerance with cryptic peptide - esp from allergen
XX or auto:antigen, esp admin orally.
XX Disclosure; Page 22-23; 38pp; English.
XX Immunological tolerance to a protein antigen may be induced by admin. of
XX a compsn. contg. at least one cryptic peptide derived from that antigen.
XX BIO mice were immunised with various Der p I peptides and then response
XX to Der p I and the peptides was measured in the presence of spleen
XX adherent cells. Peptides contg. amino acids 120-143 and 144-169 of Der p
XX I were positive i.e., they contained cryptic epitopes. The cDNA fragment
XX encoding amino acids 131-187 of Der p I was cloned into pGEX and
XX expressed in bacteria as a fusion protein with glutathione-S-transferase.
XX Mice were given orally 3 mg of this protein on 3 consecutive days, then 7
XX days later immunised subcutaneously with native Der p I in complete
XX Freund's adjuvant. Seven days later lymph nodes, stimulated with protein
XX or synthetic peptides, were assayed for lymphokines (IL-2). Mice given
XX the cryptic epitope had a much weaker response than those treated only
XX with buffer. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.le-127;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 60
Db 24 TNACISNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 83
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120
Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 143
Qy 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
Db 144 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 203
Qy 181 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPPVVIL 222
Db 204 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPPVVIL 245

RESULT 8
AAW71908
ID AAW71908 standard; protein; 245 AA.
XX
XX AAW71908;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 16-DEC-1998 (first entry)
XX
XX Dermatophagoides Der p I.
XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;
XX Der p II; Der f I; Der f II; house dust mite allergy.
XX
XX Dermatophagoides.
XX
XX US5820862-A.
XX
XX 13-OCT-1998.
XX
XX 07-JUN-1995; 95US-00482142.
XX

PR 14-APR-1993; 93WO-US003471.
PR 14-APR-1994; 94US-00227772.
PR 19-MAY-1995; 95US-00445307.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;
XX Shaked Z, Rogers BL;
XX WPI; 1998-567590/48.
XX N-PSDB; AAV61384.
XX Dermatophagoides allergen peptides - useful for treating house dust mite
XX allergy.
XX Disclosure; Col 69-72; 155pp; English.
XX The present invention describes peptides for treating sensitivity to
XX house dust mite allergens from the genus Dermatophagoides. Peptides
XX within the scope of the invention comprise at least one T cell epitope,
XX or preferably at least two T cell epitopes of a protein allergen selected
XX from the allergens Der p I, Der p II, Der f I, or Der f II. The invention
XX also describes modified peptides having similar or enhanced therapeutic
XX properties as the corresponding, naturally occurring allergen, but having
XX reduced side effects. The present sequence represents Der p I from the
XX present invention. (Updated on 25-MAR-2003 to correct PR field.) (Updated
XX on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.le-127;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 60
Db 24 TNACISNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLDLAEQ 83
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120
Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 143
Qy 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
Db 144 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 203
Qy 181 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPPVVIL 222
Db 204 QGVDDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPPVVIL 245

RESULT 9
AAW50356
ID AAW50356 standard; protein; 245 AA.
XX
XX AAW50356;
XX
XX 25-JAN-2000 (first entry)
XX
XX Dermatophagoides sp. allergen Der p I protein fragment.
XX
XX Allergen; house dust mite; detection; sensitivity; T cell epitope;
XX screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;
XX Der p I.
XX
XX Dermatophagoides sp.
XX
XX US5968526-A.
XX
XX 19-OCT-1999.
XX
XX 07-JUN-1995; 95US-00478572.
XX

```

PR 14-APR-1994; 94US-00227772.  
 PR 12-APR-1995; 95WO-US004481.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
 PI Chen X, Evans S, Kuo M;  
 XX WPI; 1999-590385/50.  
 DR N-PSDB; AA223906.  
 XX  
 XX Screening individuals for allergic reactions to T cell epitopes of major  
 PT allergens from house dust mites.  
 PT  
 XX Disclosure; Col 73-74; 158pp; English.  
 XX  
 CC This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II (DP I) and  
 CC (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the sensitivity  
 CC of the patient to house dust mites. (I) may be used to screen individuals  
 CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
 CC mite is a major cause of a variety of allergic disorders such as asthma,  
 CC rhinitis and eczematoid dermatitis. This sequence represents the house dust  
 CC mite allergen Der p I  
 XX  
 XX Sequence 245 AA;  
 SQ  
 Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 24 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGLEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGLEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKOLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPNANKIREALAQTHSAIAVIIGIKOLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDIWVRNSWDNTNGDNGYGYFAANIDLMIMTEEPYVIVIL 222  
 DB 204 QGVYDIWVRNSWDNTNGDNGYGYFAANIDLMIMTEEPYVIVIL 245  
 RESULT 10  
 AAU18959  
 ID AAU18959 standard; protein; 245 AA.  
 XX  
 AC AAU18959;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE House dust mite allergen Der p I.  
 XX  
 KW House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 PN US6268491-B1.

XX 31-JUL-2001.  
 PD  
 XX 07-JUN-1995; 95US-00484296.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 XX WPI; 2001-549074/61.  
 DR N-PSDB; AAS30721.  
 XX  
 CC Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX  
 PS Example 1; Col 73; 158pp; English.  
 XX  
 CC The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic protein from Dermatophagoides from  
 CC which the T-cell epitope containing peptides are derived  
 XX  
 SQ Sequence 245 AA;  
 Query Match 100.0%; Score 1206; DB 4; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 24 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGLEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGLEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKOLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPNANKIREALAQTHSAIAVIIGIKOLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDIWVRNSWDNTNGDNGYGYFAANIDLMIMTEEPYVIVIL 222  
 DB 204 QGVYDIWVRNSWDNTNGDNGYGYFAANIDLMIMTEEPYVIVIL 245  
 RESULT 11  
 ABG71812  
 ID ABG71812 standard; protein; 245 AA.  
 XX  
 AC ABG71812;  
 XX  
 DT 10-JAN-2003 (first entry)  
 XX  
 DE European house dust mite Der p I protein.



XX Dermatophagoides farinae protein allergen; Der f I; Der f II; allergic;  
 KW allergic response; mite allergen; house dust mite; T cell epitope;  
 KW B cell epitope; antiallergic; desensitisation therapy;  
 KW modifies allergic response of house dust mite-sensitive individual;  
 KW modify B-cell and/or T-cell response to Der f I and Der f II;  
 KW European house dust mite; Der p I.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 XX  
 XX US6423837-B1.  
 XX  
 XX 23-JUL-2002.  
 XX  
 XX 20-APR-1999; 99US-00295188.  
 XX  
 XX 13-FEB-1990; 90US-00458642.  
 XX 11-SEP-1990; 90US-00580655.  
 XX 16-AUG-1993; 93US-00107332.  
 XX 29-DEC-1993; 93US-00175071.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 XX Thomas WR, Chua K;  
 XX  
 XX WPI; 2002-672946/72.  
 XX N-PSDB; ABS56342.  
 XX  
 XX Novel isolated DNA encoding peptide from Dermatophagoides farinae protein  
 PT allergen, designated Der f I and Der f II, useful for treating and  
 PT preventing allergic responses to mite allergens, by desensitization  
 PT therapy.  
 XX  
 XX Disclosure; Fig 7; 22pp; English.  
 XX  
 XX The present invention relates to a new DNA encoding a peptide from  
 CC Dermatophagoides farinae protein allergen, designated Der f I or Der f  
 CC II, that comprises at least one epitope of the protein allergen. The  
 CC invention is useful for producing a peptide from Der f I or Der f II  
 CC protein allergen. The invention is also useful for diagnosing, treating  
 CC and preventing allergic responses to mite allergens, particularly, the  
 CC mites D. farinae. The invention is also useful as a probe for identifying  
 CC additional nucleotide sequences coding for mite allergens having amino  
 CC acid sequences similar to Der f I or Der f II. The peptide is useful for  
 CC detecting sensitivity in an individual to house dust mites and can be  
 CC used to treat sensitivity (reduce sensitivity or desensitise) in an  
 CC individual, to whom therapeutically effective quantities of D. farinae  
 CC peptide is administered. The peptides when administered to sensitive  
 CC individual modify the individual's allergic response to Der f I or Der f  
 CC II. The peptides are also useful as purified allergens useful in the  
 CC standardisation of allergen extracts or preparations which can be used as  
 CC reagents for diagnosis and treatment of allergy to house dust mites. The  
 CC epitopic peptides are useful as diagnostic reagents for determining  
 CC reactivity to the mite species. The peptides are also useful for  
 CC identifying or defining T cell epitopes and/or B cell epitopes which are  
 CC of importance in allergic responses to D. farinae allergens and to  
 CC elucidate the mediators or mechanisms of by which these reactions occur.  
 CC The present amino acid sequence represents the European house dust mite  
 CC Der p I protein as described in the invention  
 XX  
 XX Sequence 245 AA;  
 SQ  
 Query Match 100.0%; Score 1206; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 60  
 DB 24 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 143

QY 121 PPNANKIREALAQTHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPNANKIREALAQTHSAIAVIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDWIVRNSWDTNWDNGCYGYFAANIDLMIMIEEYPYVIL 222  
 DB 204 QGVYDWIVRNSWDTNWDNGCYGYFAANIDLMIMIEEYPYVIL 245  
 RESULT 12  
 AAR49920  
 ID AAR49920 standard; protein; 320 AA.  
 XX  
 XX AAR49920;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-OCT-1994 (first entry)  
 XX  
 XX Protein allergen of Der pI.  
 XX  
 XX Der pI; House Dust Mite Allergen.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 XX  
 XX Key Location/Qualifiers  
 FT Protein 1..322  
 FT /label= Der pI preproenzyme  
 FT Cleavage-site 83..85  
 FT /label= Cleavage site  
 FT /note= "proenzyme remains"  
 FT Cleavage-site 85..97  
 FT /label= cleavage site  
 FT /note= "cleavage between pro-Der pI and pre-Der pI "  
 XX  
 XX WO9405790-A1.  
 XX 17-MAR-1994.  
 XX  
 XX 10-SEP-1993; 93WO-US008518.  
 XX  
 XX 10-SEP-1992; 92US-00945288.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 XX Thomas WR, Chua K;  
 XX  
 XX WPI; 1994-101195/12.  
 XX N-PSDB; AAQ56665.  
 XX  
 XX New protein allergens of house dust mite - used for diagnosing and  
 XX treating sensitivity in an individual to house dust mite allergens.  
 XX  
 XX Example 1; Fig 21; 98pp; English.  
 XX  
 XX AAR49920 is a preproenzyme Der pI. The amino acid sequence preceding the  
 XX mature protein sequence contains cleavage sites for the pre- and  
 XX proenzyme forms, with residues 1-97 corresponding to a partial signal  
 XX peptide sequence. The mature protein can be used to detect sensitivity in  
 XX an individual to house dust mite and to reduce the sensitivity of the  
 XX individual. (Updated on 25-MAR-2003 to correct FN field.)  
 XX  
 XX Sequence 320 AA;  
 SQ  
 Query Match 100.0%; Score 1206; DB 2; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 60  
 DB 99 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 158  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218  
QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
Db 219 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 278  
QY 181 QGVYDWIVNSWDTNWDNGYGYFAANIDLMIEEYPYVUIL 222  
Db 279 QGVYDWIVNSWDTNWDNGYGYFAANIDLMIEEYPYVUIL 320

RESULT 13  
AAU07748  
ID AAU07748 standard; protein; 222 AA.  
XX  
AC AAU07748;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE House dust mite allergenic protein Der p I variant d.  
XX  
KW House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
KW T-cell epitope; polymorphic variant.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 136  
FT /note= "Wild-type Ser substituted by Thr"  
XX  
PN US6268491-B1.  
XX  
PD 31-JUL-2001.  
XX  
PF 07-JUN-1995; 95US-00484296.  
XX  
PR 16-OCT-1991; 91US-00777859.  
PR 08-MAY-1992; 92US-00881396.  
PR 14-APR-1993; 93WO-US003471.  
PR 14-APR-1994; 94US-00227772.  
PR 19-MAY-1995; 95US-00445307.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
PI Evans S, Shaked Z;  
XX  
DR WPI; 2001-549074/61.  
XX  
PT Peptides comprising T cell groups of the major allergens from  
PT Dermatophagoides (house dust mites), useful for treating house dust mite  
PT allergy in humans, and for diagnosing sensitivity to house dust mite  
PT protein allergens.  
XX  
PS Disclosure; Fig 22; 158pp; English.  
XX  
CC The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions, each  
CC region comprising at least one T cell group of a protein allergen of the  
CC genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced  
CC therapeutic properties as the naturally-occurring allergen, but have  
CC reduced side effects, and increased solubility and stability. The present  
CC sequence represents an allergenic protein from Dermatophagoides from  
CC which the T-cell epitope containing peptides are derived, a polymorphic  
CC variant of Der p I. Note: The present sequence is not shown in the

CC specification but is derived from the Der p I sequence shown in figure 22  
XX  
SQ Sequence 222 AA;  
Query Match 99.8%; Score 1203; DB 4; Length 222;  
Best Local Similarity 99.5%; Pred. No. 5.8e-127;  
Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVYDWIVNSWDTNWDNGYGYFAANIDLMIEEYPYVUIL 222  
Db 181 QGVYDWIVNSWDTNWDNGYGYFAANIDLMIEEYPYVUIL 222

RESULT 14  
AAB98347  
ID AAB98347 standard; protein; 222 AA.  
XX  
AC AAB98347;  
XX  
DT 21-AUG-2001 (first entry)  
XX  
DE D. pteronyssinus Der p 1 protein SEQ ID NO:82.  
XX  
KW Mite group 1 protein; methylotrophic yeast; Escherichia coli; allergy;  
KW recombinant mite group 1 protein; allergic response; antiallergic;  
KW infectious disease; allergic disease.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
PN WO200129078-A2.  
XX  
PD 26-APR-2001.  
XX  
PF 12-OCT-2000; 2000WO-US028204.  
XX  
PR 15-OCT-1999; 99US-0159841P.  
XX  
PA (HESK-) HESKA CORP.  
XX  
PI Best EA, McDermott MJ;  
XX  
DR WPI; 2001-308475/32.  
DR N-PSDB; AAH22385.  
XX  
PT Producing recombinant mite Group 1 protein for treating allergies.  
PT involves culturing a methylotrophic yeast microorganism or Escherichia  
PT coli transformed with nucleic acid molecule, and recovering the protein.  
XX  
PS Claim 12; Page 145; 154pp; English.  
XX  
CC The present invention describes a method for the production of a  
CC recombinant mite Group 1 protein (I). The method comprises culturing a  
CC methylotrophic yeast microorganism transformed with a nucleic acid  
CC molecule (II) encoding (I), and recovering (I), or culturing Escherichia  
CC coli transformed with (II) under conditions in which (I) forms an  
CC inclusion body in E. coli, isolating the inclusion body, and recovering  
CC (I). Also described is a method for detecting mite allergy in an animal  
CC comprising: (a) contacting (I) with a putative IgE-containing substance  
CC to form a complex between (I) and IgE, and (b) determining the presence  
CC of IgE reactive with (I) by detecting the complex, where the presence of  
CC reactive IgE is indicative of mite allergy in the animal. (I) is useful

CC for detecting mite allergy in an animal, or in a composition to reduce  
 CC allergic response to a mite Group 1 protein in a mite allergic animal.  
 CC (I) is also useful in a composition for treating or preventing allergic,  
 CC infectious or other diseases. AAH22326 to AAH22394 and AAB98326 to  
 CC AAB98349 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 XX Sequence 222 AA;

Query Match 99.5%; Score 1200; DB 4; Length 222;  
 Best Local Similarity 99.5%; Pred. No. 1.3e-126;  
 Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSTNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 1 TNACSTNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNA 180  
 DB 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNA 180  
 QY 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222  
 DB 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

RESULT 15  
 AAB98346  
 ID AAB98346 standard; protein; 302 AA.

XX AAB98346;  
 XX  
 XX 21-AUG-2001 (first entry)  
 XX  
 XX D. pteronyssinus Der p 1 protein SEQ ID NO:77.  
 XX  
 XX Mite group 1 protein; methyltrophic yeast; Escherichia coli; allergy;  
 KW recombinant mite group 1 protein; allergic response; antiallergic;  
 KW infectious disease; allergic disease.

XX Dermatophagoides pteronyssinus.

XX WO200129078-A2.

XX 26-APR-2001.

XX 12-OCT-2000; 2000WO-US028204.

XX 15-OCT-1999; 99US-0159841P.

XX (HESK-) HESKA CORP.

XX Best EA, McDermott MJ;

XX WPI; 2001-308475/32.

XX N-PSDB; AAH22381.

XX Producing recombinant mite Group 1 protein for treating allergies,  
 PT involves culturing a methyltrophic yeast microorganism or Escherichia  
 PT coli transformed with nucleic acid molecule, and recovering the protein.

XX Claim 12; Page 141-142; 154pp; English.

XX The present invention describes a method for the production of a  
 CC recombinant mite Group 1 protein (I). The method comprises culturing a  
 CC methyltrophic yeast microorganism transformed with a nucleic acid  
 CC molecule (II) encoding (I), and recovering (I), or culturing Escherichia  
 CC coli transformed with (II) under conditions in which (I) forms an  
 CC inclusion body in E. coli, isolating the inclusion body, and recovering

CC (I). Also described is a method for detecting mite allergy in an animal  
 CC comprising: (a) contacting (I) with a putative IGE-containing substance  
 CC to form a complex between (I) and IGE; and (b) determining the presence  
 CC of IGE reactive with (I) by detecting the complex, where the presence of  
 CC reactive IGE is indicative of mite allergy in the animal. (I) is useful  
 CC for detecting mite allergy in an animal, or in a composition to reduce  
 CC allergic response to a mite Group 1 protein in a mite allergic animal.  
 CC (I) is also useful in a composition for treating or preventing allergic,  
 CC infectious or other diseases. AAH22326 to AAH22394 and AAB98326 to  
 CC AAB98349 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 XX Sequence 302 AA;

Query Match 99.5%; Score 1200; DB 4; Length 302;  
 Best Local Similarity 99.5%; Pred. No. 2e-126;  
 Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSTNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 81 TNACSTNGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 140  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 141 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 200  
 QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNA 180  
 DB 201 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNA 260  
 QY 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222  
 DB 261 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 302

Search completed: May 19, 2005, 17:21:37  
 Job time : 166 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:23:37 ; Search time 132 Seconds  
(without alignments)  
562.581 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSSINGNAPAEIDLQRM.....YFAANIDLMIBEPYVYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1206	100.0	222	10	US-09-867-159A-2
2	1196	99.2	222	14	US-10-001-245-88
3	1196	99.2	320	9	US-09-877-160-2
4	1196	99.2	320	10	US-09-847-208-79
5	1196	99.2	320	14	US-10-001-245-179
6	1190	98.7	302	17	US-10-892-543-8
7	1186	98.3	303	17	US-10-892-543-32
8	1180	97.8	302	17	US-10-892-543-11
9	1176	97.5	303	17	US-10-892-543-20
10	1176	97.5	303	17	US-10-892-543-35
11	1176	97.5	303	17	US-10-892-543-38
12	1170	97.0	302	17	US-10-892-543-14
13	1166	96.7	303	17	US-10-892-543-41

14	1162	96.4	222	14	US-10-001-245-26	Sequence 26, Appl
15	1161	96.3	222	14	US-10-001-245-18	Sequence 18, Appl
16	1161	96.3	222	14	US-10-001-245-20	Sequence 20, Appl
17	1161	96.3	222	14	US-10-001-245-24	Sequence 24, Appl
18	1160	96.2	218	17	US-10-892-543-5	Sequence 5, Appl
19	1160	96.2	298	17	US-10-892-543-2	Sequence 2, Appl
20	1159	96.1	222	14	US-10-001-245-22	Sequence 22, Appl
21	1157	95.9	222	14	US-10-001-245-14	Sequence 14, Appl
22	1157	95.9	222	14	US-10-001-245-16	Sequence 16, Appl
23	1156	95.9	299	17	US-10-892-543-23	Sequence 23, Appl
24	1150	95.4	296	17	US-10-892-543-17	Sequence 17, Appl
25	1146	95.0	297	17	US-10-892-543-29	Sequence 29, Appl
26	1146	95.0	299	17	US-10-892-543-26	Sequence 26, Appl
27	1143	94.8	222	14	US-10-001-245-30	Sequence 30, Appl
28	1142	94.7	222	14	US-10-001-245-28	Sequence 28, Appl
29	1133	93.9	222	14	US-10-001-245-34	Sequence 34, Appl
30	1132	93.9	222	14	US-10-001-245-32	Sequence 32, Appl
31	1040.5	86.3	327	14	US-10-001-245-182	Sequence 182, App
32	1037.5	86.0	246	14	US-10-001-245-181	Sequence 181, App
33	1037.5	86.0	321	14	US-10-001-245-180	Sequence 180, App
34	997.5	82.7	321	10	US-09-847-208-73	Sequence 73, Appl
35	997.5	82.7	321	14	US-10-001-245-183	Sequence 183, App
36	982.5	81.5	211	10	US-09-847-208-95	Sequence 95, Appl
37	982.5	81.5	211	14	US-10-001-245-184	Sequence 184, App
38	944.5	78.3	210	14	US-10-001-245-185	Sequence 185, App
39	332.5	27.6	357	16	US-10-437-963-135411	Sequence 135411,
40	332.5	27.6	696	14	US-10-259-165-184	Sequence 184, App
41	331	27.4	363	15	US-10-425-114-48978	Sequence 48978, A
42	324.5	26.9	249	15	US-10-425-114-44401	Sequence 44401, A
43	324.5	26.9	381	15	US-10-425-114-43777	Sequence 43777, A
44	323.5	26.8	329	9	US-09-953-956-2	Sequence 2, Appl
45	323.5	26.8	329	13	US-10-114-464-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-867-159A-2  
; Sequence 2, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE OF INVENTION: and at least one anti-histamine compound  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(222)  
; OTHER INFORMATION: Peptide sequence from cystine protease.  
US-09-867-159A-2

Query Match	100.0%;	Score 1206;	DB 10;	Length 222;
Best Local Similarity	100.0%;	Pred. No. 1.3e-118;		
Matches	222;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	TNACSSINGNAPAEIDLQRMETVTPIRMQGGCGSWAFSGVAATESAYLAHNRQSLDLAEQ	60	
Db	1	TNACSSINGNAPAEIDLQRMETVTPIRMQGGCGSWAFSGVAATESAYLAHNRQSLDLAEQ	60	
Qy	61	ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYTRYVAREQSCRPNARFGISNYCQIY	120	

Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
Db 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
Db 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222

## RESULT 2

US-10-001-245-88  
; Sequence 88, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 88  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-88

Query Match 99.2%; Score 1196; DB 14; Length 222;  
Best Local Similarity 99.1%; Pred. No. 1.4e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
Db 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
Db 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222

## RESULT 3

US-09-877-160-2  
; Sequence 2, Application US/09877160  
; Publication No. US20020197268A1  
; GENERAL INFORMATION:  
; APPLICANT: Ching-Heang, Heu  
; APPLICANT: Cheng, Winston T. K.  
; TITLE OF INVENTION: ALLERGEN-CONTAINING MILK FOR ALLERGY  
; FILE REFERENCE: 12774-003001  
; CURRENT APPLICATION NUMBER: US/09/877,160  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2

; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-09-877-160-2  
Query Match 99.2%; Score 1196; DB 9; Length 320;  
Best Local Similarity 99.1%; Pred. No. 2.3e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 99 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218  
QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
Db 219 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 278  
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
Db 279 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 320

## RESULT 4

US-09-847-208-79  
; Sequence 79, Application US/09847208  
; Publication No. US20030082190A1  
; GENERAL INFORMATION:  
; APPLICANT: Saxon, Andrew  
; APPLICANT: Zhang, Ke  
; APPLICANT: Zhu, Daocheng  
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES  
; FILE REFERENCE: UC67.002A  
; CURRENT APPLICATION NUMBER: US/09/847,208  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 79  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)  
US-09-847-208-79

Query Match 99.2%; Score 1196; DB 10; Length 320;  
Best Local Similarity 99.1%; Pred. No. 2.3e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 99 TNACSGINGNAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218  
QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180  
Db 219 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 278  
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 222  
Db 279 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVVIL 320

## RESULT 5

US-10-001-245-179  
; Sequence 179, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:

APPLICANT: HOLM, Jens  
APPLICANT: IPSEN, Henrik  
APPLICANT: LARSEN, Jorgen N.  
APPLICANT: SPANGFORT, Michael D.  
TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
FILE REFERENCE: 4305/1H942-US2  
CURRENT APPLICATION NUMBER: US/10/001,245  
CURRENT FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/298,170  
PRIOR FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/249,361  
PRIOR FILING DATE: 2000-11-16  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 179  
LENGTH: 320  
TYPE: PRT  
ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-179

Query Match 99.2%; Score 1196; DB 14; Length 320;  
Best Local Similarity 99.1%; Pred. No. 2.3e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
DB 99 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158  
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120  
DB 159 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 218  
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
DB 219 PPNVKNIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 278  
QY 181 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 222  
DB 279 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 320

RESULT 6  
US-10-892-543-8  
Sequence 8, Application US/10892543  
Publication No. US20050053615A1  
GENERAL INFORMATION:  
APPLICANT: Best, Elaine A.  
APPLICANT: McDermott, Martin J.  
TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE  
FILE REFERENCE: AL-10  
CURRENT APPLICATION NUMBER: US/10/892,543  
CURRENT FILING DATE: 2004-07-15  
PRIOR APPLICATION NUMBER: 60/487,812  
PRIOR FILING DATE: 2003-07-16  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 8  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Dermatophagoides pteronyssinus  
US-10-892-543-8

Query Match 98.7%; Score 1190; DB 17; Length 302;  
Best Local Similarity 99.1%; Pred. No. 9.2e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
DB 81 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 140  
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120

DB 141 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 200  
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
DB 201 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 260  
QY 181 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 222  
DB 261 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 302

RESULT 7  
US-10-892-543-32  
Sequence 32, Application US/10892543  
Publication No. US20050053615A1  
GENERAL INFORMATION:  
APPLICANT: Best, Elaine A.  
APPLICANT: McDermott, Martin J.  
TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE  
FILE REFERENCE: AL-10  
CURRENT APPLICATION NUMBER: US/10/892,543  
CURRENT FILING DATE: 2004-07-15  
PRIOR APPLICATION NUMBER: 60/487,812  
PRIOR FILING DATE: 2003-07-16  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 32  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Dermatophagoides pteronyssinus  
US-10-892-543-32

Query Match 98.3%; Score 1186; DB 17; Length 303;  
Best Local Similarity 98.6%; Pred. No. 2.4e-116;  
Matches 219; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
DB 82 TNACSGNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 141  
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120  
DB 142 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 201  
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
DB 202 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 261  
QY 181 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 222  
DB 262 QGVYWIWRNSWDTNWDNGYGYFAANIDLMMIEEYPYVIL 303

RESULT 8  
US-10-892-543-11  
Sequence 11, Application US/10892543  
Publication No. US20050053615A1  
GENERAL INFORMATION:  
APPLICANT: Best, Elaine A.  
APPLICANT: McDermott, Martin J.  
TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE  
FILE REFERENCE: AL-10  
CURRENT APPLICATION NUMBER: US/10/892,543  
CURRENT FILING DATE: 2004-07-15  
PRIOR APPLICATION NUMBER: 60/487,812  
PRIOR FILING DATE: 2003-07-16  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 11  
LENGTH: 302  
TYPE: PRT

```
; ORGANISM: Dermatophagoides pteronyssinus
US-10-892-543-11

Query Match
Best Local Similarity 97.8%; Score 1180; DB 17; Length 302;
Matches 219; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 81 TNASSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 140
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120
DB 141 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 200
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
DB 201 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 260
QY 181 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 222
DB 261 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 302

RESULT 9
US-10-892-543-20
; Sequence 20, Application US/10892543
; Publication No. US20050053615A1
; GENERAL INFORMATION:
; APPLICANT: Best, Elaine A.
; TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE
; TITLE OF INVENTION: DUST MITE ALLERGY
; FILE REFERENCE: AL-10
; CURRENT APPLICATION NUMBER: US/10/892,543
; CURRENT FILING DATE: 2004-07-15
; PRIOR FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-892-543-20

Query Match
Best Local Similarity 97.5%; Score 1176; DB 17; Length 303;
Matches 218; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 82 TNASSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 141
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120
DB 142 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 201
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
DB 202 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 261
QY 181 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 222
DB 262 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 303

RESULT 10
US-10-892-543-35
; Sequence 35, Application US/10892543
; Publication No. US20050053615A1
; GENERAL INFORMATION:
; APPLICANT: Best, Elaine A.
; TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE
; TITLE OF INVENTION: DUST MITE ALLERGY
; FILE REFERENCE: AL-10
; CURRENT APPLICATION NUMBER: US/10/892,543
; CURRENT FILING DATE: 2004-07-15
; PRIOR FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-892-543-35

Query Match
Best Local Similarity 97.5%; Score 1176; DB 17; Length 303;
Matches 218; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 82 TNASSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 141
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120
DB 142 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 201
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
DB 202 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 261
QY 181 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 222
DB 262 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 303

RESULT 11
US-10-892-543-38
; Sequence 38, Application US/10892543
; Publication No. US20050053615A1
; GENERAL INFORMATION:
; APPLICANT: Best, Elaine A.
; TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE
; TITLE OF INVENTION: DUST MITE ALLERGY
; FILE REFERENCE: AL-10
; CURRENT APPLICATION NUMBER: US/10/892,543
; CURRENT FILING DATE: 2004-07-15
; PRIOR FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-892-543-38

Query Match
Best Local Similarity 97.5%; Score 1176; DB 17; Length 303;
Matches 218; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 82 TNASSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 141
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120
DB 142 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 201
QY 121 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180
DB 202 PPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 261
QY 181 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 222
DB 262 QGVDDYWIVRNSWDTNMGDNGYGFAANIDLMIEEYPYVVIL 303
```



Best Local Similarity 97.7%; Pred. No. 3.le-114;  
Matches 217; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TNACSIINGNAPAEIDLROMRTVTPIRMQGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 82 TNASSIINGNAPAEIDLROMRTVTPIRMQGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 141  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 61 ELVDCAHQCHGCHGTTPRGIEYIHNGVVOESYYRYVAREQSCRPNQAORFGISNYCQIY 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 142 ELVDCAHQCHGSHGDTIPRGIEYIHNGVVOESYYRYVAREQSCRPNQAORFGISNYCQIY 201  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 121 PPANKIREALAOQTHSAIAVIIGIKOLDAPRHYDGRITIIORDNGYQPNYHAVNIIVGSNA 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 202 PPNVKIREALAOQTHSAIAVIIGIKOLDAPRHYDGRITIIORDNGYQPNYHAVNIIVGSNA 261  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 181 QGVDTWIVRNWDNTNWGDNGYGYPFAANIDLMWIEEYPVVIL 222  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 262 QGVDTWIVRNWDNTNWGDNGYGYPFAANIDLMWIEEYPVVIL 303  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14  
US-10-001-245-26  
; Sequence 26, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-26

Query Match 96.4%; Score 1162; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 5.se-114;  
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSIINGNAPAEIDLROMRTVTPIRMQGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 1 TNACSIINGNAPAEIDLROMRTVTPIRMQGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 61 ELVDCAHQCHGCHGTTPRGIEYIHNGVVOESYYRYVAREQSCRPNQAORFGISNYCQIY 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 61 ELVDCAHQCHGCHGTTPRGIEYIHNGVVOESYYRYVAREQSCRPNQAORFGISNYCQIY 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 121 PPANKIREALAOQTHSAIAVIIGIKOLDAPRHYDGRITIIORDNGYQPNYHAVNIIVGSNA 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 121 PPNVKIREALAOQTHSAIAVIIGIKOLDAPRHYDGRITIIORDNGYQPNYHAVNIIVGSNA 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 181 QGVDTWIVRNWDNTNWGDNGYGYPFAANIDLMWIEEYPVVIL 222  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 181 QGVDTWIVRNWDNTNWGDNGYGYPFAANIDLMWIEEYPVVIL 222  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15  
US-10-001-245-18  
; Sequence 18, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik

; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-18

Query Match 96.3%; Score 1161; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 7e-114;  
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 TNACSSINGNAPAEIDLROMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLROMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREOSCCRPNRQRFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREOSCCRPNRQRFGISNYCQIY 120  
Qy 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180  
Qy 181 QGVYDIWVRNSWDTNMGDNGYGYFAANIDLMMIEEYPYVIVL 222  
Db 181 QGVYDIWVRNSWDTNMGDNGYGYFAANIDLMMIEEYPYVIVL 222

Search completed: May 19, 2005, 17:35:46  
Job time : 133 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:21:46 ; Search time 43 Seconds  
(without alignments)  
385.397 Million cell updates/sec

Title: US-09-867-159a-2

Perfect score: 1206

Sequence: 1 TNACSSINGNAPAEIDLQR.....YFAANIDLMIEPYVWIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1206	100.0	245	1	US-07-945-288-2
2	1206	100.0	245	1	US-08-482-831-2
3	1206	100.0	245	1	US-08-461-809-2
4	1206	100.0	245	1	US-08-461-441-2
5	1206	100.0	245	2	US-08-482-142-2
6	1206	100.0	245	2	US-08-478-572-2
7	1206	100.0	245	3	US-08-484-296-2
8	1206	100.0	245	5	PCT-US93-08518-2
9	1206	100.0	320	1	US-07-945-288-10
10	1206	100.0	320	1	US-08-462-831-10
11	1206	100.0	320	1	US-08-461-809-10
12	1206	100.0	320	1	US-08-461-441-10
13	1206	100.0	320	5	PCT-US93-08518-10
14	1177	97.6	222	1	US-07-945-288-11
15	1177	97.6	222	1	US-08-462-831-11
16	1177	97.6	222	1	US-08-461-809-11
17	1177	97.6	222	1	US-08-461-441-11
18	1177	97.6	222	5	PCT-US93-08518-11
19	1144	94.9	245	3	US-08-460-040-2
20	997.5	82.7	321	1	US-07-945-288-6
21	997.5	82.7	321	1	US-08-462-831-6
22	997.5	82.7	321	1	US-08-461-809-6
23	997.5	82.7	321	1	US-08-461-441-6
24	997.5	82.7	321	2	US-08-482-142-6
25	997.5	82.7	321	2	US-08-478-572-6
26	997.5	82.7	321	3	US-08-484-296-6
27	997.5	82.7	321	5	PCT-US93-08518-6

Sequence 195, App  
Sequence 195, App  
Sequence 195, App  
Sequence 197, App  
Sequence 197, App  
Sequence 197, App  
Sequence 197, App  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 20, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 1, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 36, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-07-945-288-2  
; Sequence 2, Application US/07945288  
; Patent No. 5433948  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Wayne R.  
; APPLICANT: Chua, Kaw-Yan  
; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/945,288  
; FILING DATE: 19920910  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 580,655  
; FILING DATE: 11 SEPTEMBER 1990  
; APPLICATION NUMBER: 458,642  
; FILING DATE: 13 FEBRUARY 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRAGOURAS, AMY E.  
; REGISTRATION NUMBER: P36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-945-288-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.4e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 24 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
QY 61 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 120  
Db 84 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 143  
QY 121 PPANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
Db 144 PPANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMIEEYPYVVIL 222  
Db 204 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMIEEYPYVVIL 245

## RESULT 2

US-08-462-831-2  
; Sequence 2, Application US/08462831  
; Patent No. 5552142  
; GENERAL INFORMATION:

; APPLICANT:  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,831  
; FILING DATE:

; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/945,288  
; FILING DATE: 10 SEPTEMBER 1992  
; APPLICATION NUMBER: US 580,655  
; FILING DATE: 11 SEPTEMBER 1990  
; APPLICATION NUMBER: US 458,642  
; FILING DATE: 13 FEBRUARY 1990

; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRAGOURAS, AMY E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-462-831-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.4e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 24 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
QY 61 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 120

Db 84 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 143  
QY 121 PPANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
Db 144 PPANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMIEEYPYVVIL 222  
Db 204 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMIEEYPYVVIL 245

## RESULT 3

US-08-461-809-2  
; Sequence 2, Application US/08461809  
; Patent No. 5770202  
; GENERAL INFORMATION:

; APPLICANT:  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,809  
; FILING DATE:

; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/945,288  
; FILING DATE: 10 SEPTEMBER 1992  
; APPLICATION NUMBER: US 580,655  
; FILING DATE: 11 SEPTEMBER 1990  
; APPLICATION NUMBER: US 458,642  
; FILING DATE: 13 FEBRUARY 1990

; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRAGOURAS, AMY E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-461-809-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.4e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 24 TNACSNAGAPAEIDLQRMRTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 83  
QY 61 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 120  
Db 84 ELVDCASQHGCHGDTTPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAQRFGISNYCQIY 143  
QY 121 PPANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180



RESULT 6  
US-08-478-572-2  
; Sequence 2, Application US/08478572  
; Patent No. 5988526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-June-1995  
; APPLICATION NUMBER: US/08/478,572  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-478-572-2

Query Match 100.0%; Score 1206; DB 2; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.4e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
DB 24 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 83  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143  
QY 121 PPANKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
DB 144 PPANKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
DB 204 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 7  
US-08-484-296-2

; Sequence 2, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,296  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 245 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-484-296-2

Query Match 100.0%; Score 1206; DB 3; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.4e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
DB 24 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 83  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143  
QY 121 PPANKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
DB 144 PPANKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
DB 204 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 8  
PCT-US93-08518-2  
; Sequence 2, Application PC/TUS9308518  
; GENERAL INFORMATION:

```
;
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08518
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: F36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-08518-2

Query Match 100.0%; Score 1206; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.4e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQMRITVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 24 TNACSSINGNAPAEIDLQMRITVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120
Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 143
Qy 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180
Db 144 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 203
Qy 181 QGVDTYIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222
Db 204 QGVDTYIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 9
US-07-945-288-10
; Sequence 10, Application US/07945288
; Patent No. 5433948
; GENERAL INFORMATION:
; APPLICANT: Thomas, Wayne R.
; APPLICANT: Chua, Kaw-Yan
; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
```

```
;
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,288
; FILING DATE: 19920910
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: F36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-945-288-10

Query Match 100.0%; Score 1206; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQMRITVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 99 TNACSSINGNAPAEIDLQMRITVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 158
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCOIY 218
Qy 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180
Db 219 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 278
Qy 181 QGVDTYIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222
Db 279 QGVDTYIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 320

RESULT 10
US-08-462-831-10
; Sequence 10, Application US/08462831
; Patent No. 5552142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-462-831-10

Query Match 100.0%; Score 1206; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLROMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 99 TNACSGNAPAEIDLROMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGLSNYCIY 120
DB 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGLSNYCIY 218
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180
DB 219 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 278
QY 181 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222
DB 279 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 320

RESULT 11
US-08-461-809-10
; Sequence 10, Application US/08461809
; Patent No. 5770202
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,809
; FILING DATE:
; CLASSIFICATION: 424
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-809-10

Query Match 100.0%; Score 1206; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLROMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 99 TNACSGNAPAEIDLROMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGLSNYCIY 120
DB 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGLSNYCIY 218
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180
DB 219 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 278
QY 181 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 222
DB 279 QGVDDYWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVVIL 320

RESULT 12
US-08-461-441-10
; Sequence 10, Application US/08461441
; Patent No. 5773002
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,441
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
```



```
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-08518-10
;
Query Match 100.0%; Score 1206; DB:1; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60
Db 99 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 158
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQRFGISNYCOIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQRFGISNYCOIY 218
Qy 121 PPNAKIREALAOATHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 219 PPNAKIREALAOATHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278
Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIMIEEYPVWVIL 222
Db 279 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIMIEEYPVWVIL 320

RESULT 13
PCT-US93-08518-10
; Sequence 10, Application PC/TUS9308518
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08518
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-08518-10
;
Query Match 100.0%; Score 1206; DB:5; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60
Db 99 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 158
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQRFGISNYCOIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQRFGISNYCOIY 218
Qy 121 PPNAKIREALAOATHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 219 PPNAKIREALAOATHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278
Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIMIEEYPVWVIL 222
Db 279 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMIMIEEYPVWVIL 320

RESULT 14
US-07-945-288-11
; Sequence 11, Application US/07945288
; Patent No. 5433948
; GENERAL INFORMATION:
; APPLICANT: Thomas, Wayne R.
; APPLICANT: Chua, Kaw-Yan
; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,288
; FILING DATE: 19920910
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: P36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
```

```
/ NAME/KEY: misc feature
/ LOCATION: 50
/ OTHER INFORMATION: /label=Xaa is His or Tyr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 81
/ OTHER INFORMATION: /label=Xaa is Glu or Lys
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 124
/ OTHER INFORMATION: /label=Xaa is Ala or Val
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 136
/ OTHER INFORMATION: /label=Xaa is Ser or Thr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 215
/ OTHER INFORMATION: /label=Xaa is Glu or Gln
US-07-945-288-11

Query Match          97.6%; Score 1177; DB 1; Length 222;
Best Local Similarity 97.7%; Pred. No. 2.5e-126;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
Db 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120

Qy 121 PPNANKIREALQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180
Db 121 PPNANKIREALQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

Qy 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIEEYPYVVIL 222
Db 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIEEYPYVVIL 222
```

```
RESULT 15
US-08-462-831-11
; Sequence 11, Application US/08462831
; Patent No. 5552142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
```

```
/ FILING DATE: 13 FEBRUARY 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MANDRAGOURAS, AMY E.
/ REGISTRATION NUMBER: 36,207
/ REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 227-5941
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 222 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 50
/ OTHER INFORMATION: /label=Xaa is His or Tyr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 81
/ OTHER INFORMATION: /label=Xaa is Glu or Lys
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 124
/ OTHER INFORMATION: /label=Xaa is Ala or Val
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 136
/ OTHER INFORMATION: /label=Xaa is Ser or Thr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 215
/ OTHER INFORMATION: /label=Xaa is Glu or Gln
US-08-462-831-11
```

```
Query Match          97.6%; Score 1177; DB 1; Length 222;
Best Local Similarity 97.7%; Pred. No. 2.5e-126;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
Db 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120

Qy 121 PPNANKIREALQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180
Db 121 PPNANKIREALQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

Qy 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIEEYPYVVIL 222
Db 181 QGVDDYWIVRNSWDTNWDNGYGYFAANIDLMIEEYPYVVIL 222
```

Search completed: May 19, 2005, 17:32:32  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:39:08 ; Search time 39 Seconds  
(without alignments)  
547.695 Million cell updates/sec

Title: US-09-867-159A-2  
Perfect score: 1206  
Sequence: 1 TNACSSINGNAPAEIDLQRMR.....YFAANIDLMWIEYPVVIL 222

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 114359

Minimum DB seq length: 0  
Maximum DB seq length: 222

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	982.5	81.5	211	2 S21864	probable cysteine
2	389.5	32.3	94	2 S03380	major fecal allerg
3	315	26.1	215	2 A59428	ervatamin B (EC 3.
4	311	25.8	221	2 A59041	cysteine proteinas
5	301	25.0	216	2 S06837	glycyl endopeptida
6	298	24.7	218	1 KHCYL	cathepsin L (EC 3.
7	295	24.5	221	2 A59040	cysteine proteinas
8	284.5	23.6	217	2 S15844	cathepsin S (EC 3.
9	274	22.7	214	2 S46476	cysteine proteinas
10	273	22.6	218	2 S67481	cathepsin L-like c
11	240	19.9	212	2 S03964	stem bromelain (EC
12	209	17.3	183	2 A4938	cysteine proteinas
13	207	17.2	184	2 S02729	actinidain (EC 3.4
14	185.5	15.4	166	2 B4938	cysteine proteinas
15	179	14.8	95	2 PQ0650	senescence-associa
16	171	14.2	165	2 C44938	cysteine proteinas
17	156.5	13.0	139	2 A41404	cathepsin L (EC 3.
18	145	12.0	174	2 B48454	cathepsin B-like c
19	139	11.5	155	2 S57426	cysteine proteinas
20	129.5	10.7	110	2 A61061	actinidain (EC 3.4
21	129.5	10.7	157	2 S57451	cysteine proteinas
22	125.5	10.4	136	2 S57624	cysteine proteinas
23	123.5	10.2	150	2 S57425	cysteine proteinas
24	121	10.0	152	2 T25581	hypothetical prote
25	120.5	10.0	152	2 S57423	cysteine proteinas
26	116.5	9.7	152	2 S57421	cysteine proteinas
27	115	9.5	145	2 S60456	cysteine proteinas
28	111.5	9.2	30	2 B27634	major fecal allerg
29	111	9.2	152	2 S57422	cysteine proteinas

ALIGNMENTS

RESULT 1

S21864  
Probable cysteine proteinase (EC 3.4.22.-) - Euroglyphus maynei  
N;Alternate names: allergen Eur m i  
C;Species: Euroglyphus maynei  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: S21864  
R;Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.  
submitted to the EMBL Data Library, June 1991  
A;Reference number: S21864  
A;Accession: S21864  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-211 <KEN>  
A;Cross-references: UNIPROT:P25780; EMBL:X60073  
C;Genetics:  
A;Introns: 100/3; 155/2  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase

Query Match	81.5%;	Score 982.5;	DB 2;	Length 211;
Best Local Similarity	83.9%;	Pred. No. 4.3e-81;		
Matches	177;	Conservative	17;	Mismatches 16; Indels 1; Gaps 1;
Qy	1	TNACSSING-NAPAEIDLQRMTVTPIRMQCGSCWAFSGVAATESAYLAHRNQLDLAE	59	
Db	1	TYACSSINSVSLPSELDLRLTPTPIRMQCGSCWAFSGVASTESAYLAHRNQLDLAE	60	
Qy	60	QELVDCASQHGCHGDTIPRGIEYIQNGVVOESYRYRVAREQSCRRPNARFGISNYCOI	119	
Db	61	QELVDCASQHGCHGDTIPRGIEYIQNGVVOEHYYPVAREQSCRRPNARQRYGLKNYCOI	120	
Qy	120	YPNANKIRREALAQTSAIAVIIGIKDLDAFRHVDGRTTIQRDNGYQPNYHAVNIVGYSN	179	
Db	121	SPDPSNKIRALQTHTTAVAVIIGIKDLNAPRHYDGRITMQHDNGYQPNYHAVNIVGYN	180	
Qy	180	AQGVYDVIVRNSWDTTWGDNNGYGYFAANIDL 210		
Db	181	TQGVYDVIVRNSWDTTWGDNNGYGYFAANINL 211		

RESULT 2

S03380  
major fecal allergen Der p 1 - house-dust mite (Dermatophagoides pteronyssinus) (fragmen  
N;Alternate names: allergen Der p1  
C;Species: Dermatophagoides pteronyssinus  
C;Date: 05-Mar-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S03380  
R;Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.  
Protein Seq. Data Anal. 2, 17-21, 1989  
A;Title: Structural studies on the allergen Der p1 from the house dust mite Dermatophago  
A;Reference number: A31657; MUID:89098855; PMID:2911558



Db 178 KKYWIVKNSGKKGDKGYIYMAKDRKXHCIGIATAASPLV 218

RESULT 7

A59040

Cysteine proteinase I (EC 3.4.22.-) - ginger

C:Species: Zingiber officinale (ginger)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C:Accession: A59040

R;Laursen, R.A.

submitted to the Protein Sequence Database, July 1999

A:Description: Amino acid sequences of cysteine proteases from ginger rhizome, Zingiber

A:Reference number: A59040

A:Accession: A59040

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-221 <LAU>

A:Cross-references: UNIPROT:P82473

A:Experimental source: rhizome

A:Note: Residues 80-87 were not determined but are based on mass measurement and similar

C:Superfamily: papain

C:Keywords: cysteine proteinase; glycoprotein; hydrolase

F:24-65,58-98,155-206/Disulfide bonds: #status predicted

F:27,161,181/Active site: Cys, His, Asn #status predicted

F:95,156/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.5%; Score 295; DB 2; Length 221;

Best Local Similarity 32.0%; Pred. No. 3.8e-19;

Matches 65; Conservative 35; Mismatches 91; Indels 12; Gaps 5;

Qy 11 PAEIDLQMRVTVPINQGGCGSCWAFSGVAATESAYLAHRNQLDLAQELVDCASQ-H 69

Db 4 PDSIDWREKGAVPVKNQGGCGSCWAFDAIAAVEGINQIVTGDLSLSBQQLVDCSTRNH 63

Qy 70 GCGGDTTPRGLEYIOHN-GVVOESYRVRVAREQSC-RRENAQRFGSNYCQIYPPNANKI 127

Db 64 GCGGWPYAFQIYINNGINSEHYPTGTNGTCDTKEHAHVSDSRNVPNSDEKSL 123

Qy 128 REALAQTHSAIAVIGIKDLDAFRH--YDGRITIIQRDNGYQPNYHAVNIVGYSNAQQVDY 185

Db 124 QKAVANQPVSVTMDAAGRDFQLYRNGIFTGSCNISAN-----HVRTVGGRETENDKDY 176

Qy 186 WIVRNSWDTNWDNGYGYFAANI 208

Db 177 WTVKNSWGNKNGESGYIRVERNI 199

RESULT 8

S15844

cathepsin S (EC 3.4.22.27) - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jul-2004

C:Accession: S15844; S23680; S16972; S23957

R;Ritonja, A.; Colic, A.; Dolenc, I.; Ogrinc, T.; Podobnik, M.; Turk, V.

FEBS Lett. 283, 329-331, 1991

A:Title: The complete amino acid sequence of bovine cathepsin S and a partial sequence o

A:Reference number: S15844; MUID:91257334; PMID:2044774

A:Accession: S15844

A:Molecule type: protein

A:Residues: 1-217 <RIT>

A:Cross-references: UNIPROT:P25326

R;Wiederanders, B.; Broemme, D.; Kirschke, H.; Kalkkinen, N.; Rinne, A.; Paquette, T.; T

FEBS Lett. 286, 189-192, 1991

A:Title: Primary structure of bovine cathepsin S. Comparison to cathepsins L, H, B and p

A:Reference number: S16972; MUID:91323515; PMID:1864368

A:Accession: S23680

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 22-217 <WIE>

A:Cross-references: GB:M95211; NID:g162814; PIDN:AAA30435.1; PID:g162815

A:Note: 143-Pro was also found

A:Accession: S16972

A:Molecule type: protein



```
Qy 127 IRELAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNAQGVYD 186
Db 121 MMYAVSQPIITVAV-----DANANFOYKSGVFNPGCTSLN-HAVTAIGY----GQDSI 170
Qy 187 IVRNSWDITNMGDNGYGYFAANI 208
Db 171 IYPKWGAKEAGYIIRWARDV 192

RESULT 12
A44938
cysteine proteinase (EC 3.4.22.-) - Trypanosoma cruzi (fragment)
C;Species: Trypanosoma cruzi
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A44938
R;Eakin, A.E.; Bouvier, J.; Sakanari, J.A.; Craik, C.S.; McKerrow, J.H.
Mol. Biochem. Parasitol. 39, 1-8, 1990
A;Title: Amplification and sequencing of genomic DNA fragments encoding cysteine protease
A;Reference number: A44938; MUID:90158686; PMID:2406590
A;Accession: A44938
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: DNA
A;Residues: 1-183 <EAK>
A;Cross-references: UNIPROT:Q9GPN3; UNIPROT:Q8T2Y4; UNIPROT:Q26883; UNIPROT:Q26884; GB:M
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

Query Match 17.3%; Score 209; DB 2; Length 183;
Best Local Similarity 29.9%; Pred. No. 1.7e-11;
Matches 60; Conservative 23; Mismatches 82; Indels 36; Gaps 7;

Qy 10 APABIDLRQMTVTPIRMQCGCGSWAFSGVAAATESAYLAHRNOSLDLAEOELVDC-ASQ 68
Db 1 APAADVWRKAGVAVTKDQCGCGSWAFSAIGNVSGWQFLAGHPLTNLSEQLVSCDKTD 60
Qy 69 HGCHGDTIPRGIEYI---QHNVVQESYRYRYVAREQSCRPNRQAORFGISNYCQIYPPNAN 125
Db 61 SGCSGGLMNAFVWQENNGVYTEDSYPSASGE-----GISPPCTTSHTVG 109
Qy 126 KI---REALAQTHSAIAVIIGIKDLDAFRH-----YDGRITII-----QRDNGYQPNVHA 171
Db 110 ATITGHVELPDEQAIAWLAVNGPVAHASSWMTYTGVTSCVSEQLD-----HG 162
Qy 172 VNIVGYSNAQGVYDVIWRNSW 192
Db 163 LLLVGYNDSAAVPYVIVKNSW 183

RESULT 13
S02729
actinidin (EC 3.4.22.14) precursor (clone pAC.7) - kiwi fruit (fragment)
C;Species: Actinidia chinensis (kiwi fruit)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 22-Jun-1999
C;Accession: S02729
R;Praekelt, U.M.; McKee, R.A.; Smith, H.
Plant Mol. Biol. 10, 193-202, 1988
A;Title: Molecular analysis of actinidin, the cysteine proteinase of Actinidia chinensis
A;Reference number: S02728
A;Accession: S02729
A;Molecule type: mRNA
A;Residues: 1-184 <PRA>
A;Cross-references: EMBL:X13139; NID:g15958; PIDN:CAA31529.1; PID:g15959
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
F;1-159/Product: actinidin (fragment) #status experimental <MAT>

Query Match 17.2%; Score 207; DB 2; Length 184;
Best Local Similarity 34.2%; Pred. No. 2.5e-11;
Matches 51; Conservative 19; Mismatches 57; Indels 22; Gaps 6;

Qy 70 GCHGDTIPRGIEYIQHNQVQ--ESYRYVAREQSCR--RPNARQFGISNYCQIYPPNANK 126
Db 3 CGNGGYTDGQFIINNGINTENYPTAODGECNLDLQNEKYVTIDTVENV--PYNNE 60
```

```
Qy 127 IRELAQTHSAIAVIIGIKDLDAFRHYDG-----RTIIQRDNGYQPNYHVNIVGYSN 179
Db 61 WALQTAVTYQPSVALDAAG-DAFKHYSSGIFTGPGCTAID-----HAVTIVGVT 110
Qy 180 AQGVYDVIWRNSWDITNMGDNGYGYFAANI 208
Db 111 EGGIDYVIVKNSWDITNMGEEGYMRLRN 139
```

## RESULT 14

```
B44938
cysteine proteinase (EC 3.4.22.-) - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B44938
R;Eakin, A.E.; Bouvier, J.; Sakanari, J.A.; Craik, C.S.; McKerrow, J.H.
Mol. Biochem. Parasitol. 39, 1-8, 1990
A;Title: Amplification and sequencing of genomic DNA fragments encoding cysteine protease
A;Reference number: A44938; MUID:90158686; PMID:2406590
A;Accession: B44938
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: DNA
A;Residues: 1-166 <EAK>
A;Cross-references: UNIPROT:Q26742; GB:M27306
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

Query Match 15.4%; Score 185.5; DB 2; Length 166;
Best Local Similarity 33.5%; Pred. No. 1.9e-09;
Matches 58; Conservative 25; Mismatches 75; Indels 15; Gaps 8;

Qy 28 QGCGCGSWAFSGVAAATESAYLAHRNOSLDLAEOELVDCASQHGCHGDTIPRGIEYI-QHN 86
Db 1 QGCGCGSWAFSTIGNIEGQVAGNPLVLSLSEQILVYCDPLCGCGGLMDNFAFWIVNSN 60
Qy 87 G--VWQESYRYVA--REQSCRPNRQAORFG--ISNYCQIYPPNANKIREALAQTHS-AIA 139
Db 61 GGNVFTASYPYVSGNGEQPCQCMNGHEIGAATIDHVDL-PQDEDAIAAVALAENRPLAIA 119
Qy 140 VIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNAQGVYDVIWRNSW 192
Db 120 V-----EAPFYGHNGGYIILTCTSEQLD-HGVLLVGYNDNSNPPYVIVKNSW 166
```

## RESULT 15

```
PQ0650
senescence-associated protein SAG2 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: PQ0650
R;Hensel, L.L.; Grbic, V.; Baumgarten, D.A.; Bleecker, A.B.
Plant Cell 5, 553-564, 1993
A;Title: Developmental and age-related processes that influence the longevity and senesc
A;Reference number: PQ0650; MUID:93299122; PMID:8518555
A;Accession: PQ0650
A;Molecule type: mRNA
A;Residues: 1-95 <HEN>
A;Cross-references: UNIPROT:Q9LL83
C;Comment: This protein is a senescence-associated protein.
C;Genetics:
A;Gene: SAG2
C;Superfamily: papain

Query Match 14.8%; Score 179; DB 2; Length 95;
Best Local Similarity 42.2%; Pred. No. 3.8e-09;
Matches 38; Conservative 15; Mismatches 33; Indels 4; Gaps 2;

Qy 11 PAEIDLRQMTVTPIRMQCGCGSWAFSGVAAATESAYLAHRNOSLDLAEOELVDC---S 67
Db 6 PETKDWREDGIVSPVKDQCGCGSWFTSTTGALEAAVHQAFKGISLSEQQLVDCAGAFN 65
Qy 68 QHGCHGDTIPRGIEYIQHNQVQ--ESYRY 96
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:35:53 ; Search time 175 Seconds  
(without alignments)  
649.609 Million cell updates/sec

Title: US-09-867-159A-2  
Perfect score: 1206  
Sequence: 1 TNACSSINGNAPAEIDLQRMR.....YFAANIDLMMEIBPYVIVL 222

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 737793

Minimum DB seq length: 0  
Maximum DB seq length: 222

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03: +  
1: uniprot\_sprot: +  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	944.5	78.3	210	2 Q9GY0	Q9GY0 dermatophag
2	607	50.3	146	2 Q95X05	Q95X05 dermatophag
3	469	38.9	107	2 Q95X04	Q95X04 dermatophag
4	437	36.2	133	2 Q968Y3	Q968Y3 dermatophag
5	389.5	32.3	94	2 Q7M431	Q7M431 dermatophag
6	382	31.7	221	2 Q95P34	Q95P34 blomia trop
7	315	26.1	215	1 ERVB_TABDI	P60994 tabernaemon
8	313.5	26.0	219	2 Q9NGW1	Q9NGW1 fasciola gi
9	311	25.8	221	1 GPII_ZINOF	P82474 zingiber of
10	300	24.9	218	1 CATL_CHICK	P09648 gallus gall
11	295	24.5	213	1 MDOL_PSEMR	P83443 pseudananas
12	295	24.5	221	1 GPI_ZINOF	P82473 zingiber of
13	291	24.1	214	2 QWSH4	QWSH4 ancylostoma
14	286	23.7	208	1 ERVC_TABDI	P83654 tabernaemon
15	284.5	23.6	217	1 CATS_BOVIN	P25326 bos taurus
16	279	23.1	178	2 Q9SDN1	Q9SDN1 prunus dulc
17	276	22.9	215	2 Q6TXV9	Q6TXV9 glycine max
18	276	22.9	216	2 Q26559	Q26559 spirometra
19	275.5	22.8	219	2 Q9NGW0	Q9NGW0 fasciola gi
20	267	22.1	189	2 Q966T7	Q966T7 spirometra
21	266.5	22.1	181	2 Q9NH98	Q9NH98 stylonychia
22	265.5	22.0	189	2 Q8T8B8	Q8T8B8 halichondri
23	258	22.0	196	2 Q8I889	Q8I889 leptinotars
24	257	21.3	222	2 Q868H3	Q868H3 leishmania
25	253.5	21.0	166	2 Q24943	Q24943 fasciola he
26	249.5	20.7	217	1 CATL_SHEEP	Q10991 ovis aries
27	248.5	20.6	166	2 Q24942	Q24942 fasciola he
28	241	20.0	196	2 Q8I888	Q8I888 leptinotars
29	240	19.9	212	1 BROM_ANACO	P14518 ananas como
30	239.5	19.9	166	2 Q24945	Q24945 fasciola he
31	238.5	19.8	189	2 Q9NHV3	Q9NHV3 babesia equ

32	237.5	19.7	176	2	Q94C46	Q94C46 carica cand
33	236.5	19.6	167	2	Q8I881	Q8I881 leptinotars
34	232.5	19.3	181	2	Q84XA1	Q84XA1 carica cand
35	231	19.2	217	2	Q9U0C5	Q9U0C5 clonorchis
36	230	19.1	166	2	Q6LAF8	Q6LAF8 homo sapien
37	228.5	18.9	191	2	Q6JZ26	Q6JZ26 fundulus he
38	227.5	18.9	177	2	Q70B20	Q70B20 platichthys
39	226.5	18.8	166	2	Q24948	Q24948 fasciola he
40	223	18.5	175	2	Q6LAF7	Q6LAF7 homo sapien
41	222.5	18.4	168	2	Q43947	Q43947 sarcocystis
42	220	18.2	174	2	Q868H2	Q868H2 leishmania
43	219.5	18.2	163	2	Q9SPU9	Q9SPU9 hordeum vul
44	218.5	18.1	166	2	Q24946	Q24946 fasciola he
45	214	17.7	165	2	Q8T4J2	Q8T4J2 pagumogonim

ALIGNMENTS

RESULT 1	
ID	Q9GY0
AC	PRELIMINARY;
DT	PRT; 210 AA.
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Allergen Der fi (Fragment).
OS	Dermatophagoides farinae (House-dust mite).
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC	Acariiformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC	Pyroglyphidae; Dermatophagoides.
OX	NCBI_TaxID=6954;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Hao M.Q., Xu J., Zhong N.S.;
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: belongs to peptidase family C1.
DR	EMBL; AF285763; AAG00520.1; -.
DR	PIR; A27634; A27634.
DR	HSP; P80067; 1JQP.
DR	GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR	GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR	InterPro; IPR000668; Peptidase_C1.
DR	Pfam; PF00112; Peptidase_C1; 1.
DR	PRINTS; PR00705; PAPAIN_.
DR	SMART; SM00645; Pept C1; 1.
DR	PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
DR	PROSITE; PS00139; THIOL_PROTEASE CYS; 1.
DR	PROSITE; PS00639; THIOL_PROTEASE HIS; UNKNOWN_1.
KW	Hydrolase; Protease; Thiol protease.
FT	NON TER 1
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON TER 210
FT	NON

```

Db      181 QGVYWIVRNWDTTWGDGCGYGFQAGNNL 210
RESULT 2
Q95X05 PRELIMINARY; PRT; 146 AA.
AC Q95X05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cysteine proteinase (Fragment).
GN Name=CPW2;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194431; AAL14424.1; -.
DR HSSP; P43235; 1BY8.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
FT NON_TER 1
FT NON_TER 146
SQ SEQUENCE 146 AA; 16852 MW; BB304800946D4047 CRC64;

Query Match 50.3%; Score 607; DB 2; Length 146;
Best Local Similarity 84.2%; Pred. No. 1.5e-46;
Matches 112; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 39 GVAATESAYLAHRNOSLDLAQEQLVDCASQHGCHGDTIPRGIEYIQHNGVQSYRYVA 98
Db      1 GVAATESAYLAHRNOSLDLAQEQLVDCASQHGCHGDTIPRGIEYIQHNGVQSYRYVA 60
QY 99 REQSCRPNNAQRGINSYQIYPPNANKIREALQTHSAIAVIGIKDLDAFRHYDGRIT 158
Db      61 REQSCRPNNSQHYGINSYQIYPPDVKQIREALTQTHTAIAVIGIKDLRAFOHYDGRIT 120
QY 159 IQRDNGYQPNYHA 171
Db      121 IQHDNGYQPNYHA 133

RESULT 3
Q95X04 PRELIMINARY; PRT; 107 AA.
AC Q95X04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cysteine proteinase (Fragment).
GN Name=CPW3;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194432; AAL14425.1; -.
DR InterPro; IPR000169; Pept cys acsite.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 12277 MW; A80E7876CBA6F97A CRC64;

Query Match 38.9%; Score 469; DB 2; Length 107;
Best Local Similarity 80.4%; Pred. No. 2.5e-34;
Matches 86; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 77 PRGIEYIQHNGVQSYRYVAREQSCRPNNAQRFGISNYCQIYPPNANKIREALQTHS 136
Db      1 PRGIEYIQHNGVQSYRYVAREQSCRPNNSQHYGINSYQIYPPDVKQIREALTQTHT 60
QY 137 ATAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAINVIGYSNAQGV 183
Db      61 ATAVIIGIKDLRAFOHYDGRITIIQHDNGYQPNYHAINVIGYSTQGV 107

RESULT 4
Q98Y3 PRELIMINARY; PRT; 133 AA.
AC Q98Y3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cysteine proteinase (Fragment).
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE FROM N.A.
RA Park H., Yun H.C., Kim K.Y., Park S.Y., Park S.K.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF145247; AAK38773.1; -.
DR HSSP; P60994; 1IWD.
DR MEROPS; C01.073; -.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
FT NON_TER 1
FT NON_TER 133
SQ SEQUENCE 133 AA; 14965 MW; 5033C26B15E68E9C CRC64;

Query Match 36.2%; Score 437; DB 2; Length 133;
Best Local Similarity 61.4%; Pred. No. 2.4e-31;
Matches 81; Conservative 17; Mismatches 34; Indels 0; Gaps 0;

QY 31 CGSCWAFSGVAATESAYLAHRNOSLDLAQEQLVDCASQHGCHGDTIPRGIEYIQHNGVQ 90
Db      1 CGSCWAFAGVAATESAYLAHRNOSLDLAQEQLVDCASQHGCHGDTIPRGIDYIQNGIVE 60
QY 91 ESYRYVAREQSCRPNNAQRFGISNYCQIYPPNANKIREALQTHSAIAVIGIKDLDAF 150
Db      61 EQAYEYNARENCEPPEPRHSIEQYQIDHSNVELIKTALDKYKSAVAVINIHNINAF 120
QY 151 RHYDGRITIIQRD 162
Db      121 RHYDGSYVITTD 132

RESULT 5
Q7M431 PRELIMINARY; PRT; 94 AA.
AC Q7M431;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major fecal allergen Der p I (Fragments).
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE.

```

```

RX MEDLINE=89098855; PubMed=2911558;
RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;
RT "Structural studies on the allergen Der p1 from the house dust mite
RT Dermatophagoides pteronyssinus: similarity with cysteine
RL proteinases.";
DR PIR; S03380; S03380.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR ProDom; PD000158; Peptidase_C1; 1.
FT NON_TER 1 94
FT SEQUENCE 94 AA; 10327 MW; 9BF744165C8428A8 CRC64;

Query Match 32.3%; Score 389.5; DB 2; Length 94;
Best Local Similarity 48.6%; Pred. No. 2.9e-27;
Matches 88; Conservative 0; Mismatches 4; Indels 89; Gaps 4;

QY 1 TNACISNGNAPAEIDLQMRVTPIR--MQGGCGSCWAFSGVAATESAYLAHRNQLDLA 58
Db 1 TNACISNGNAPAEIDLQMRVTPIRQMGGCGSXXAFSGVA----- 43
QY 59 EQELVDCASQHGCHDTPRGIEYIOHNGVVQESYRYVAREQSCRPNAPRGISNYCQ 118
Db 44 -----GIEYIOHNGVVQESY-----RFGISNYCQ 68
QY 119 IYPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITTIQRDNGYQPNVHAVNIGYS 178
Db 69 IYPNANK-----DNGYQPNVHAVNIGYX 93
QY 179 N 179
Db 94 N 94

RESULT 6
ID Q95PJ4 PRELIMINARY; PRT; 221 AA.
AC Q95PJ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Cysteine protease (Fragment).
OS Blomia tropicalis (Mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea;
OC Echinopodidae; Blomia.
OX NCBI_TaxID=40697;
RN [1]
RP SEQUENCE FROM N.A.
RA Mora C.I., Diaz A.M., Montealegre F., Flores I.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF277840; AAK58415.1; -.
DR HSP; P53634; IK3B.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Protease.
FT NON_TER 1 221
FT CHAIN <1 221 cysteine protease.
SQ SEQUENCE 221 AA; 25126 MW; 272B45EA53F2900 CRC64;

Query Match 31.7%; Score 382; DB 2; Length 221;
Best Local Similarity 38.4%; Pred. No. 3.7e-26;
Matches 86; Conservative 35; Mismatches 77; Indels 26; Gaps 8;

```

```

QY 11 PAEIDLQMRVTPIRQMGGCGSCWAFSGVAATESAYLAHRNQLDLAEOELVDCASQ-- 68
Db 2 PANFWRQKTHVNPPIRNMGGCGSCWAFSGVAATESAYLAHRNQLDLAEOELVDCASQ-- 61
QY 69 -----HGCHGDTPIRGIEYIOHNGVVQESYRYVAREQSCRPNAPRGISNYCQ 120
Db 62 DPTVKHCGSCGMSPEAFKTMKQKLESHYPMKLNQC-QANVGRTHVSY----- 116
QY 121 PPNANKIREALQTHSAI-----AVIIGIKDLA-FRHYDGRITTIQRDNGYQPNV--HAVN 173
Db 117 --NSLRYRAGDQEIQAIRAMNHGPPVVIYHGTAEHFRNL--RKGLRGAGYNDQAIDHAVV 172
QY 174 IVGYSNAQGVYVIRNSWNTWGDNGYGYFAANIDLMIEEYP 217
Db 173 LVGWTQNGIDYIVRTISWGTOMGDAGYGFVERHNSLGINNY 216

RESULT 7
ERVB TABDI STANDARD; PRT; 215 AA.
AC P60994;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ervatom B (EC 3.4.22.-) (ERV-B).
OS Tabernaemontana divaricata (Crepe jasmine) (Ervatamia coronaria).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Gentianales; Apocynaceae; Rauvolfioideae; Tabernmontantaneae;
OC Tabernaemontana.
OX NCBI_TaxID=52861;
RN [1]
RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA Biswas S., Chakrabarti C., Kundu S., Jagannadham M.V.,
RA Dattagupta J.K.;
RT "Proposed amino acid sequence and the 1.63 A X-ray crystal structure
RT of a plant cysteine protease, ervatamin B: some insights into the
RT structural basis of its stability and substrate specificity.";
RL Proteins 51:489-497(2003).
RN [2]
RP SEQUENCE OF 1-21, AND CHARACTERIZATION.
RC TISSUE=Latex;
RX PubMed=10691612; DOI=10.1021/jf990661j;
RA Kundu S., Sundd M., Jagannadham M.V.;
RT "Purification and characterization of a stable cysteine protease
RT ervatamin B, with two disulfide bridges, from the latex of Ervatamia
RT coronaria.";
RL J. Agric. Food Chem. 48:171-179(2000).
CC -!- FUNCTION: Cysteine protease.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Laticifer.
CC -!- PTM: Not glycosylated.
CC -!- MISCELLANEOUS: Active over the pH range 3.0-10.5 and up to 62
CC degrees Celsius. Stable in 8 M urea and 2.5 M GuHCl at neutral pH,
CC in 40% acetonitrile, 70% ethanol and 50% methanol. Unstable in
CC SDS.
CC -!- SIMILARITY: Belongs to the peptidase C1 family.
DR PIR; A59428; A59428.
DR PDB; 1LWD; X-ray; A=1-215.
DR MEROPS; C01.099; -.
DR ProDom; PD000158; Peptidase C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; FALSE NEG.
KW 3D-structure; Direct protein sequencing; Hydrolase; Thiol protease.
FT ACT_SITE 25 25
FT ACT_SITE 158 158 By similarity.
FT ACT_SITE 178 178
FT DISULFID 22 63
FT DISULFID 56 96
FT DISULFID 152 203
SQ SEQUENCE 215 AA; 23183 MW; 4DE62E43BA4F4F83 CRC64;

```

```

Query Match      26.1%; Score 315; DB 1; Length 215;
Best Local Similarity 34.5%; Pred. No. 3.7e-20;
Matches 76; Conservative 32; Mismatches 90; Indels 22; Gaps 7;

QY 11 PABIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQLDLAQELVDC-ASQH 69
Db 2 PSFVDRSKGAVNSIKNQKCGSCWAFSAVAESINKRTGQLISLSEQLVDCDTASH 61
QY 70 GCHGDTIPRGIEVIQNGVVO-BSYRYVAREOSCRPNQAQREGISNYCQIYPPNANKIR 128
Db 62 GCGNGMNAFQYIITNGGIDTQONTYPYSAVQSGCKPYRLRVVVSINGFORVTRNNEALQ 121
QY 129 EALAQTHSALAVIIGIKDLDA----FRHYDGRTHIQRDNGYQPNYHVNIVGVYNAQGV 184
Db 122 SAVASQVSVTV-----EACAPQHYSS-GIFTGCGTAQN-HGVVIVGYGTOSGKN 172
QY 185 YWTVRNSWDTNMGDNGYGYFAANI-----DLMMIEEYP 217
Db 173 YWTVRNSWGNQGNQGYIMWERNVASSAGLCGTAQLPSYP 212

RESULT 8
Q9NGW1 PRELIMINARY; PRT; 219 AA.
AC Q9NGW1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cathepsin L (Fragment).
GN Names-cat-L1E;
OS Fasciola gigantica (Giant liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
OX NCBI_TaxID=46835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21332182; PubMed=11438432; DOI=10.1016/S1383-5769(01)00068-X;
RA Grams R., Vichasari-Grams S., Sobhon P., Upatham E.S., Viyanant V.;
RT "Molecular cloning and characterization of cathepsin L encoding genes
from Fasciola gigantica.";
RL Parasitol. Int. 50:105-114(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Grams S.V., Grams R., Sobhon P., Viyanant V., Upatham E.S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family C1.
DR EMBL; AF239267; AAP44678.1; -.
DR HSSP; P53634; 1K3B.
DR MEROPS; C01.033; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOI_PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE HIS; 1.
KW Hydrolase; Protease; Thiol protease.
FT NON_TER 1
SQ SEQUENCE 219 AA; 24217 MW; 7F443104B4071D3C CRC64;

Query Match      26.0%; Score 313.5; DB 2; Length 219;
Best Local Similarity 32.4%; Pred. No. 5.2e-20;
Matches 73; Conservative 39; Mismatches 88; Indels 25; Gaps 7;

QY 11 PABIDLRQMTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQLDLAQELVDC-AS 67
Db 2 PDKIDRESGYTEVDKQKCGSCWAFSTTGTWEGQTMKNRTSISFSEQLVDCSGPWG 61
QY 68 QHGHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQREGISNYCQIYPPNA 124

```

```

Db 62 NYGCMGLMENAYEYLKQFLETSESYPTAVEDQCR--YNRLGVAKYTDYTVHSGSE 119
QY 125 NKIREAL-AQTHSATAVIIGIKDLDA-FRHYDGRTHIQRDNGYQPNYHVNIVGVYNAQ 182
Db 120 VELKNLVGAEGPAAVAV-----DVSEDFMYSGGIYQSRSTCSLRVNHAVLAIVGYGTGG 174
QY 183 VDYWTVRNSWDTNMGDNGYGYFAAN-----IDLMMIEEYP 217
Db 175 TDYWTVRNSWGSWGERGYIRVVRNKGNGMGSIASLASLPVAFPP 219

RESULT 9
GP11_ZINOF
ID GP11_ZINOF STANDARD; PRT; 221 AA.
AC P82474;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cysteine proteinase GP-II (EC 3.4.22.-).
OS Zingiber officinale (Ginger).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Zingiberaceae;
OC Zingiber.
OX NCBI_TaxID=94328;
RN [1]
RP SEQUENCE.
TS TISSUE=Root;
RX MEDLINE=20156257; PubMed=10691991;
RA Choi K.H., Laursen R.A.;
RT "Amino-acid sequence and glycan structures of cysteine proteases with
proline specificity from ginger rhizome Zingiber officinale.";
RL Eur. J. Biochem. 267:1516-1526(2000).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage with a proline residue
at P2.
CC -!- SIMILARITY: Belongs to the peptidase C1 family.
CC PIR; A59041; A59041.
DR PDB; 1COD; X-ray; A/B/C/D=1-221.
DR MEROPS; C01.017; -.
DR GlycoSuiteDB; P82474; -.
DR InterPro; IPR000169; Pept_cys_acsite.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; FALSE_NEG.
DR PROSITE; PS00139; THIOI_PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE HIS; 1.
KW 3D-structure; Direct protein sequencing; Glycoprotein; Hydrolase;
Thiol protease.
FT ACT_SITE 27 27 By similarity.
FT ACT_SITE 161 161 By similarity.
FT DISULFID 24 65 By similarity.
FT DISULFID 58 98 By similarity.
FT DISULFID 155 206 By similarity.
FT CARBOHYD 99 99 N-linked (GlcNAc...)/FTId-CAR_000190.
FT CARBOHYD 156 156 N-linked (GlcNAc...)/FTId-CAR_000200.
FT STRAND 7 8
FT TURN 9 13
FT STRAND 20 20
FT TURN 22 23
FT STRAND 25 25
FT HELIX 27 44
FT STRAND 50 50
FT HELIX 52 58
FT TURN 60 61
FT TURN 64 65
FT STRAND 66 66
FT TURN 70 80
FT TURN 81 81

```

```

FT STRAND 83 83
FT STRAND 85 85
FT TURN 86 88
FT STRAND 107 107
FT STRAND 111 114
FT STRAND 120 127
FT TURN 128 129
FT STRAND 132 136
FT HELIX 141 144
FT TURN 145 145
FT STRAND 150 151
FT STRAND 161 171
FT TURN 172 173
FT STRAND 174 180
FT STRAND 183 183
FT TURN 185 186
FT TURN 188 188
FT STRAND 189 189
FT TURN 190 191
FT STRAND 192 196
FT TURN 202 203
FT HELIX 205 207
FT TURN 208 209
FT STRAND 213 216
SQ SEQUENCE 221 AA; 23922 MW; 909A312BD8632D42 CRC64;

Query Match 25.8%; Score 311; DB 1; Length 221;
Best Local Similarity 33.8%; Pred. No. 8.8e-20;
Matches 69; Conservative 34; Mismatches 89; Indels 12; Gaps 5;

QY 11 PAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQLVDC-ASQH 69
DB 4 PDSIDWRENGAVPVKNQGGCGSCWAFSTVAAVEGINQIVTGDLISLSEQLVDCITANH 63

QY 70 GCHGDTIPRGIEYIOHN-GVVOESYRYRYVAREQSCRRP-NAQRFGISNYCOIYPPNANKI 127
DB 64 GCRGGWNPAPQFIVNNGGINSERTYPRGQDGLCNSTVNAPVVSIDSYNVFSHNEQSL 123

QY 128 REALAQTSAIAVIGIKLDAPFRH--YDGRITIQRDNGQPNYHVNIVGYSNAQGVY 185
DB 124 QKAVANQPVSTMDAAGRDQLYRSGLTSCNISAN-----HALTVVGYGTENDKDF 176

QY 186 WIVRNSWDTWNGDNGYGFANID 209
DB 177 WIVKNSWGNWGESGYRAERNIE 200

RESULT 10
CATL_CHICK STANDARD; PRT; 218 AA.
AC P09648;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cathepsin L (EC 3.4.22.15) (Fragments).
GN Name=CTSL;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=87304227; PubMed=3305012;
RA Wada K., Takai T., Tanabe T.;
RT "Amino acid sequence of chicken liver cathepsin L.";
RL Eur. J. Biochem. 167:13-18(1987).
RN [2]
RP SEQUENCE OF 1-37 AND 177-216.
RX MEDLINE=87080783; PubMed=3792553; DOI=10.1016/0014-5793(86)81137-1;
RA Wada K., Tanabe T.;
RT "N-terminal amino acid sequences of the heavy and light chains of

```

```

RT chicken liver cathepsin L.";
RL FEBS Lett. 209:330-334(1986).
RN [3]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=88050863; PubMed=3676277;
RA Dufour E., Obled A., Valin C., Bechet D., Ribadeau-Dumas B.,
RA Huet J.-C.;
RT "Purification and amino acid sequence of chicken liver cathepsin L.";
RL Biochemistry 26:5689-5695(1987).
CC -I- FUNCTION: Important for the overall degradation of proteins in
CC lysosomes.
CC -I- CATALYTIC ACTIVITY: Specificity close to that of papain. As
CC compared to cathepsin B, cathepsin L exhibits higher activity
CC towards protein substrates, but has little activity on Z-Arg-Arg-
CC NHMeC, and no peptidyl-dipeptidase activity.
CC -I- SUBUNIT: Dimer of a heavy and a light chain linked by disulfide
CC bonds.
CC -I- SUBCELLULAR LOCATION: Lysosomal.
CC -I- SIMILARITY: Belongs to the peptidase C1 family.
DR HSSP; P07711; ICJL.
DR MEROPS; C01.032; -.
DR InterPro; IPR000169; Pept_cys_acsite.
DR InterPro; IPR000668; Peptidase_C1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SM000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Direct protein sequencing; Glycoprotein; Hydrolase; Lysosome;
KW Thiol protease.
FT CHAIN 1 176 Cathepsin L heavy chain.
FT NON CONS 176 177
FT CHAIN 177 218 Cathepsin L light chain.
FT ACT_SITE 25 25 By similarity.
FT ACT_SITE 165 165 By similarity.
FT ACT_SITE 185 185 By similarity.
FT CARBOHYD 109 109 N-linked (GlcNAc...).
FT DISULFID 22 65 By similarity.
FT DISULFID 56 99 By similarity.
FT DISULFID 158 207 Interchain (By similarity).
FT CONFLICT 21 21 Q -> I (in Ref. 3).
FT CONFLICT 29 29 S -> N (in Ref. 3).
FT CONFLICT 40 42 RTK -> FKT (in Ref. 3).
FT CONFLICT 97 98 Missing (in Ref. 3).
FT CONFLICT 109 109 N -> K (in Ref. 3).
FT CONFLICT 177 177 Missing (in Ref. 3).
FT CONFLICT 197 197 I -> Q (in Ref. 3).
SQ SEQUENCE 218 AA; 23963 MW; D26E21BC071686C2 CRC64;

Query Match 24.9%; Score 300; DB 1; Length 218;
Best Local Similarity 31.7%; Pred. No. 8.4e-19;
Matches 70; Conservative 37; Mismatches 100; Indels 14; Gaps 7;

QY 10 APAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQLVDC--- 66
DB 1 APRSDWREKGYVTPVKDQGGCGSCWAFSTTGALGEGHFKTKGLVSLSEQLVDCSRPE 60

QY 67 SQHCCHDITPRGIEYIOHN-GVVOESYRYRYVAR-EQSCR-RPNAQRFGISNYCOIYPPN 123
DB 61 GNQCGGLMDQAFQVYVDNGGIDSESYPTAKDDECKRYKAEYNAANDTFVDIPOGH 120

QY 124 ANKIREALAQTHS-AIAVIGIKDLDAFRHYDGRITIQRDNGQPNYHVNIVGYSNAQ 182
DB 121 ERLMKAVASVGPVSVAIDAG---HSSFQYQSGIYEPDCSSEDLDHGLVVGVEGG 177

QY 183 VDYWVRNSWDTWNGDNGYGFAA-----NIDLMMIEEYPVY 219
DB 178 KTYWIVKNSWGEKWDGKIYMAKDRKNHCGIATAASYPLV 218

RESULT 11

```

Db	63	CKGGWNRAYDFIISNNVTTDENTPYRAYQTC---	NANYFPNSAYITGSYYVRNDESL	111
Qy	128	REALAQTSAIAVIIGIKDADFRRHYDGRTHIIQ	RNDGNGYOPNYHVAIVNIGYNAQGVYWI	187
Db	120	HMMYAVSNQPIAALIDASG-DNFQYK-GVSG	PGFSLN-HAITIIGYGRD---SYWI	173
Qy	188	VRNSWDTNWGDNGY	201	
Db	174	VRNSGSSWGQGY	187	
RESULT 12				
GPI_ZINOF				
ID	GPI_ZINOF	STANDARD;	PRT;	221 AA.
AC	P82473;			
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	25-OCT-2004	(Rel. 45, Last annotation update)		
DE	Cysteine proteinase Gp-I	(EC 3.4.22.-)		
DE	Submitted	(AUG-2002) to Swiss-Prot.		
OC	Eukaryota; Viridiplantae; Streptophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida;			
OC	Zingiber.			
OC	NCBI_TaxID=94328;			
OX	[1]			
RN	SEQUENCE.			
RP	TISSUE=Root;			
RC	MEDLINE=20156257; PubMed=10691991;			
RX	Choi K.H., Laursen R.A.;			
RA	"Amino-acid sequence and glycan structures of cysteine proteases with			
RT	proline specificity from ginger rhizome Zingiber officinale.";			
RL	Eur. J. Biochem. 267:1516-1526(2000).			
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage with a proline residue			
CC	at P2.			
CC	-I- SIMILARITY: Belongs to the peptidase C1 family.			
CC	-I- CAUTION: The authors regard the sequence as tentative, as they			
CC	believe that it may have been contaminated by a homologous			
CC	protein.			
DR	PIR: A59040; A59040.			
DR	HSSP: P82474; LCQD.			
DR	MEROFS; C01.017; -.			
DR	InterPro; IPR000169; Pept_Cys_acsite.			
DR	Pfam; PF00112; Peptidase_C1.			
DR	PRINTS; PR00705; PAPAIN_1.			
DR	ProDom; PD000158; Peptidase_C1; 1.			
DR	SMART; SM00645; Pept_C1; 1.			
DR	PROSITE; PS00640; THIOI_PROTEASE ASN; 1.			
DR	PROSITE; PS00139; THIOI_PROTEASE CYS; FALSE NEG.			
DR	PROSITE; PS00639; THIOI_PROTEASE HIS; FALSE NEG.			
KW	Direct protein sequencing; Glycoprotein; Hydrolase; Thiol protease.			
FT	ACT_SITE 27 27			
FT	ACT_SITE 161 161			
FT	DISULFID 24 65			
FT	DISULFID 58 98			
FT	DISULFID 155 206			
FT	CARBOHYD 95 95			
FT	CARBOHYD 156 156			
FT	UNSURE 2			
SQ	SEQUENCE 221 AA; 24241 MW; 3035D7870EA743DB CRC64;			
Query Match 24.5%; Score 295; DB 1; Length 221;				
Best Local Similarity 32.0%; Pred. No. 2.4e-18;				
Matches 65; Conservative 35; Mismatches 91; Indels 12; Gaps 5;				
Qy	11	PAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLDLAEQELVDCASQ-H	69	
Db	4	PDSIDWREKGA VVPVKNGCGCGSWAFDAIAAVEGINQIVTGDLSLSEQLVDCSTRNH	63	
Qy	70	GCHGDTIPRGIEYIQHN-GVVQESYRYRYVAREQSC-RPNAQRGINSYCCYIYPNANKI	127	
Db	64	GCEGGWPYRAFOYIINNGGINSSEHYPYTGNTGTCTKENAHVVSIDSRYRVPSNDEKSL	123	

Qy 128 REALAQTHSAIAVIGIKDLDAFRH--YDGRITIIQRDNGVQPNTHAVNIQVSNAGVDY 185  
 Db 124 QKAVANQFVSVTMDAAGRCDFOLYRNGIFGTSNCSNAN-----HYRTVGGRETENDKDY 176  
 Qy 186 WIVRNSWDTNWDGNGYGVFAANI 208  
 Db 177 WTVKNSGKNWGESGYRVERNI 199

RESULT 13  
 Q8WSH4 ID Q8WSH4 PRELIMINARY; PRT; 214 AA.  
 AC Q8WSH4; 2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Cathepsin L-like protease (Fragment).  
 GN Name=cpl-1;  
 OS Ancylostoma caninum (Dog hookworm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
 OX NCBI\_TaxID=29170;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22072558; PubMed=12076767; DOI=10.1016/S0166-6851(02)00066-X;  
 RA Britton C., Murray L.;  
 RT "A cathepsin L protease essential for Caenorhabditis elegans  
 RT embryogenesis is functionally conserved in parasitic nematodes.";  
 RL Mol. Biochem. Parasitol. 122:21-33(2002).  
 CC -I- SIMILARITY: Belongs to peptidase family C1.  
 DR EMBL; AF320084; AAU37181.1; -.  
 DR HSSP; PF0711; 1CJL.  
 DR MEROPS; C01.066; -.  
 DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPAIN.  
 DR SMART; SM00645; Pept C1; 1.  
 DR PROSITE; PS00133; CARBOXYPEPT ZN 2; UNKNOWN 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE HIS; 1.  
 KW Hydrolase; Protease; Thiol protease.  
 FT NON TER 1  
 SQ SEQUENCE 214 AA; 23486 MW; 4A339A1A7A61E525 CRC64;

Query Match 24.1%; Score 291; DB 2; Length 214;  
 Best Local Similarity 33.3%; Pred. No. 5.3e-18;  
 Matches 68; Conservative 35; Mismatches 85; Indels 16; Gaps 7;

Qy 14 IDLRQMTVTPIRQGGCGSCWAFSGVAATESAYLAHRNOSLDLAQELVDCSQ---HG 70  
 Db 1 VQWRDKGLVTEVKNQCGSCWAFSATGALEGHARASQGVSLSEQLVDCSTKYGNHG 60  
 Qy 71 CHGDTIPRGIEYIOHN-GVQVESYRVRVAREQSCRRAORFGISN-YQIYPPNANKIR 128  
 Db 61 CNGGLMDLAFEYIKDNRHGDITSESYVYGRDMKCHKFKKDKTGAVDNGVDLPEDGEALX 120  
 Qy 129 EALA-QTHSAIAVIGIKDLDAFR--HYDGRITIIQRDNGVQPNTHAVNIQV-SNAQGV 183  
 Db 121 IAVATQGFISAIDAGHRTFQLYKKGYYVD-----ECSSELDHGLVLLVGYGTDPDPEAG 174  
 Qy 184 DYWIVRNSWDTNWDGNGYGVFAAN 207  
 Db 175 DYWLKNSWGTGCKGKYIRARN 198

RESULT 14  
 ERVC TABDI ID ERVC TABDI STANDARD; PRT; 208 AA.  
 AC P83654; 2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ervatin C (EC 3.4.22.-) (ERV-C).  
 OS Tabernaemontana divaricata (Crepe Jasmine) (Ervatamia coronaria).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Gentianales; Apocynaceae; Rauvolfioideae; Tabernaemontaneae;  
 OC Tabernaemontana.  
 OX NCBI\_TaxID=52861;  
 RN [1]  
 RP SEQUENCE, X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS), MASS SPECTROMETRY,  
 AND DISULFIDE BONDS  
 RX PubMed=14769029; DOI=10.1021/bi0357659;  
 RA Guha Thakurta P., Biswas S., Chakrabarti C., Sundd M.,  
 RA Jagannadham M.V., Dattagupta J.K.;  
 RT "Structural basis of the unusual stability and substrate specificity  
 RT of ervatamin C, a plant cysteine protease from Ervatamia coronaria";  
 RL Biochemistry 43:1532-1540(2004).  
 RN [2]  
 RP SEQUENCE, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RC TISSUE=Latex;  
 RA Guha Thakurta P., Biswas S., Chakrabarti C., Dattagupta J.K.;  
 RL Submitted (AUG-2003) to Swiss-Prot.  
 RN [3]  
 RP SEQUENCE OF 1-21, FUNCTION, AND ENZYME REGULATION.  
 RC TISSUE=Latex;  
 RX PubMed=9836431;  
 RA Sundd M., Kundu S., Pal G.P., Medicherla J.V.;  
 RT "Purification and characterization of a highly stable cysteine  
 RT protease from the latex of Ervatamia coronaria";  
 RL Biosci. Biotechnol. Biochem. 62:1947-1955(1998).  
 RN [4]  
 RP STABILITY.  
 RC TISSUE=Latex;  
 RX PubMed=10543984; DOI=10.1006/bbrc.1999.1550;  
 RA Kundu S., Sundd M., Jagannadham M.V.;  
 RT "Structural characterization of a highly stable cysteine protease  
 RT ervatamin C";  
 RL Biochem. Biophys. Res. Commun. 264:635-642(1999).  
 CC -I- FUNCTION: Cysteine proteinase. Hydrolyzes denatured natural  
 CC substrates such as casein, hemoglobin, azoalbumin and azocasein  
 CC with a high specific activity. Has little or no activity against  
 CC synthetic substrates.  
 CC -I- ENZYME REGULATION: Activated by thio-specific activators such as  
 CC cysteine, beta-mercaptoethanol, DTT and glutathione. Inhibited by  
 CC the thiol-specific inhibitors leupeptin, iodoacetamide, PCMB, NEM,  
 CC mercuric chloride and sodium tetrathionate.  
 CC -I- SUBUNIT: Monomer.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- TISSUE SPECIFICITY: Laticifer.  
 CC -I- MASS SPECTROMETRY: MW=23000; METHOD=MALDI; RANGE=1-208;  
 CC NOTE=Ref.1.  
 CC -I- MISCELLANEOUS: The enzyme has a Km of 9.09 microm. Optimum  
 CC temperature for enzyme activity is 50 degrees Celsius, and optimum  
 CC pH is between 7.5 and 8.0. The enzyme is very stable and retains  
 CC its structure and activity up to 70 degrees Celsius and over a  
 CC wide pH range (between 2.0 and 12.0).  
 CC -I- SIMILARITY: Belongs to the peptidase C1 family.  
 DR PDB; 1O0B; X-ray; -.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR PRINTS; PR00705; PAPAIN.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SM00645; Pept C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE HIS; FALSE NEG.  
 KW 3D-structure; Direct protein sequencing; Hydrolase; Thiol protease.  
 FT ACT SITE 25 25 By similarity.  
 FT ACT SITE 157 157 By similarity.  
 FT ACT SITE 173 173 By similarity.  
 FT DISULFID 22 63  
 FT DISULFID 56 96  
 FT DISULFID 114 193





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:32:37 ; Search time 161 Seconds  
(without alignments)  
533.297 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSGINGNAPAEIDLQRM.....YFAANIDLMIEEYPPVVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 1481396

Minimum DB seq length: 0  
Maximum DB seq length: 222

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206	100.0	222	5	Aao20568 Cysteine
2	1206	100.0	222	5	Abb98533 Cysteine
3	1203	99.8	222	4	Aau07748 House dus
4	1200	99.5	222	4	Aab98347 D. pteron
5	1196	99.2	222	4	Aau07746 House dus
6	1196	99.2	222	5	Abg67023 House dus
7	1196	99.2	222	8	Adk52140 Der pl al
8	1196	99.2	222	8	Adr87225 Dust mite
9	1189	98.6	222	4	Aau07747 House dus
10	1177	97.6	222	2	Aar52742 Protein a
11	1162	96.4	222	5	Abg67024 House dus
12	1162	96.4	222	5	Abg67030 House dus
13	1161	96.3	222	5	Abg67027 House dus
14	1161	96.3	222	5	Abg67029 House dus
15	1160	96.2	222	5	Abg67026 House dus
16	1159	96.1	222	5	Abg67028 House dus
17	1157	95.9	222	5	Abg67025 House dus
18	1143	94.8	222	5	Abg67032 House dus
19	1142	94.7	222	5	Abg67031 House dus
20	1139	94.4	211	2	Aay25678 Euroglyph
21	1139	94.4	211	7	Adc34926 Euroglyph
22	1133	93.9	222	5	Abg67034 House dus
23	1132	93.9	222	5	Abg67033 House dus
24	982.5	81.5	211	2	Aay25677 Euroglyph
25	982.5	81.5	211	2	Aay25676 Euroglyph

26	982.5	81.5	211	7	ADC34925	Adc34925 Euroglyph
27	982.5	81.5	211	7	ADC34924	Adc34924 Euroglyph
28	949.5	78.7	212	2	AAY25679	Aay25679 Euroglyph
29	949.5	78.7	212	7	ADC34927	Adc34927 Euroglyph
30	472	39.1	86	8	ADR87224	Adr87224 Dust mite
31	472	39.1	159	8	ADR87234	Adr87234 Chimeric
32	460	38.1	86	8	ADR87223	Adr87223 Dust mite
33	430	35.7	81	8	ADR87222	Adr87222 Dust mite
34	430	35.7	153	8	ADR87233	Adr87233 Chimeric
35	407	33.7	72	4	AAU07749	Aau07749 House dus
36	372.5	30.9	181	2	AAW72342	Aaw72342 Dermatoph
37	372.5	30.9	181	2	AAV50544	Aav50544 Dermatoph
38	372.5	30.9	181	4	AAU19147	Aau19147 Synthetic
39	354.5	29.4	181	2	AAW72343	Aaw72343 Dermatoph
40	354.5	29.4	181	2	AAV50545	Aav50545 Dermatoph
41	354.5	29.4	181	4	AAU19148	Aau19148 Synthetic
42	308.5	25.6	211	2	AAW44779	Aaw44779 Human cat
43	303	25.1	217	8	ADL92138	Adl92138 Chymopapa
44	301.5	25.0	190	3	AAV83106	Aav83106 DI4059 pr
45	295	24.5	215	5	AAE28367	Aae28367 Dirofilar

ALIGNMENTS

RESULT 1  
AAO20568  
ID AAO20568 standard; protein; 222 AA.  
XX AAO20568;  
AC AAO20568;  
XX 02-JAN-2003 (first entry)  
DT 02-JAN-2003 (first entry)  
XX Cysteine protease protein.  
DE Cysteine protease protein.  
XX Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema.  
XX Dermatophagoides pteronyssinus.  
OS Dermatophagoides pteronyssinus.  
XX Key Location/Qualifiers  
FH Misc-difference 105  
FT /note= "Encoded by ACC"  
XX WO200278736-A2.  
XX 10-OCT-2002.  
PD 28-MAR-2002; 2002WO-FR001098.  
XX 30-MAR-2001; 2001FR-00004370.  
PR 03-MAY-2001; 2001FR-00005929.  
PR 29-MAY-2001; 2001US-00867159.  
XX (ANTI-) ANTIALIS SARR.  
XX Loria E, Terrasse G, Trehin Y;  
XX WPI; 2002-750636/81.  
XX N-PSDB; AAL41281.  
XX Antiallergic compositions containing an anti-histamine, a histamine  
PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
PT for the allergen.  
XX Claim 13; Page 30-31; 32pp; French.  
XX The invention relates to antiallergic compositions containing an anti-  
CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
CC isolated nucleic acid molecule that has at least one polynucleotide  
CC sequence coding for the allergen, together with a pharmaceutical carrier.

CC The pharmaceutical composition of the invention is useful as a non-  
 CC specific anti-allergic treatment, and also useful in the treatment of  
 CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
 CC allergic and atopic eczema. This sequence represents the cysteine  
 CC proenzyme protein relating to the anti-allergic compositions of the  
 CC invention  
 XX  
 XX Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSJDLAEQ 60  
 DB 1 TNACISNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSJDLAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYVAREQSCRPNQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYVAREQSCRPNQRFGISNYCQIY 120  
 QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYNSA 180  
 DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYNSA 180  
 QY 181 QGVVDYWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPVWIL 222  
 DB 181 QGVVDYWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPVWIL 222

## RESULT 2

ABB98533  
 ID ABB98533 standard; protein; 222 AA.  
 AC ABB98533;  
 DT 13-DEC-2002 (first entry)  
 DE Cysteine protease.  
 XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunotherapy; allergen; allergic hypersensitivity reaction;  
 KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
 KW cysteine protease.  
 XX Dermatophagoides pteronyssinus.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 105  
 FT /note= "Encoded by CCA"  
 XX  
 XX FR2822709-A1.  
 XX  
 XX PD 04-OCT-2002.  
 XX  
 XX PF 03-MAY-2001; 2001FR-00005929.  
 XX  
 XX PR 30-MAR-2001; 2001FR-00004370.  
 XX  
 XX PA (ANTI-) ANTIALIS SARI.  
 XX  
 XX PI Loria E, Terrasse G, Trehin Y;  
 XX  
 XX WPI; 2002-735037/80.  
 DR N-PSDB; ABQ80833.  
 XX  
 XX Antiallergic composition, useful for preventing and treating e.g. asthma,  
 PT rhinitis or eczema, containing at least two of allergen, antihistamine  
 PT and histamine synthesis inhibitor.  
 XX  
 XX PS Claim 7; Page 27-28; 33pp; French.  
 XX  
 CC The present invention relates to an antiallergic pharmaceutical

CC composition (I) comprising a pharmaceutical carrier containing an active  
 CC agent combination of at least two of: an allergen; an antihistamine; and  
 CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
 CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
 CC rhinitis or allergic atopic eczema, in babies, children or adults. The  
 CC present sequence is cysteine protease from Dermatophagoides  
 CC pteronyssinus, which was used as an allergen in the invention  
 XX  
 XX Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSJDLAEQ 60  
 DB 1 TNACISNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSJDLAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYVAREQSCRPNQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYVAREQSCRPNQRFGISNYCQIY 120  
 QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYNSA 180  
 DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYNSA 180  
 QY 181 QGVVDYWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPVWIL 222  
 DB 181 QGVVDYWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPVWIL 222

## RESULT 3

AAU07748  
 ID AAU07748 standard; protein; 222 AA.  
 XX  
 AC AAU07748;  
 DT 04-DEC-2001 (first entry)  
 DE House dust mite allergenic protein Der p I variant d.  
 XX  
 XX House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope; polymorphic variant.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 136  
 FT /note= "Wild-type Ser substituted by Thr"  
 XX  
 XX US6268491-B1.  
 XX  
 XX PD 31-JUL-2001.  
 XX  
 XX PF 07-JUN-1995; 95US-00484296.  
 XX  
 XX PR 16-OCT-1991; 91US-00777859.  
 XX PR 08-MAY-1992; 92US-00881396.  
 XX PR 14-APR-1993; 93WO-US003471.  
 XX PR 14-APR-1994; 94US-00227772.  
 XX PR 19-MAY-1995; 95US-00445307.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 XX  
 XX WPI; 2001-549074/61.  
 XX  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite

PT protein allergens.

XX Disclosure; Fig 22; 158pp; English.

XX The invention relates to an isolated peptide of the major protein

XX allergens of the genus Dermatophagoides, which comprises at least one T

CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I

CC or DF II. The isolated peptide comprises at least two regions, each

CC region comprising at least one T cell group of a protein allergen of the

CC genus Dermatophagoides. The regions are derived from the same or

CC different protein allergens of the genus Dermatophagoides. The peptides

CC are useful for treating house dust mite allergy in humans. The peptides

CC are also useful for detecting or diagnosing sensitivity to house dust

CC mite protein allergens. The present peptides have similar or enhanced

CC therapeutic properties as the naturally-occurring allergen, but have

CC reduced side effects, and increased solubility and stability. The present

CC sequence represents an allergenic protein from Dermatophagoides from

CC which the T-cell epitope containing peptides are derived, a polymorphic

CC variant of Der p I. Note: The present sequence is not shown in the

CC specification but is derived from the Der p I sequence shown in figure 22

XX

SQ Sequence 222 AA;

Query Match 99.8%; Score 1203; DB 4; Length 222;

Best Local Similarity 99.5%; Pred. No. 5.8e-127;

Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQMRVTVPTRMQGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60

DB 1 TNACSGNAPAEIDLQMRVTVPTRMQGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQSSYYRYVAREQSCRRPNAQRFGISNYCOIY 120

DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQSSYYRYVAREQSCRRPNAQRFGISNYCOIY 120

QY 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180

DB 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180

QY 181 QGVVDYWIVRNSWDNTNWDNGNGYGYFAANIDLMWIEEYPVWIL 222

DB 181 QGVVDYWIVRNSWDNTNWDNGNGYGYFAANIDLMWIEEYPVWIL 222

RESULT 4

AAB98347

ID AAB98347 standard; protein; 222 AA.

XX AAB98347;

XX 21-AUG-2001 (first entry)

XX D. pteronyssinus Der p 1 protein SEQ ID NO:82.

DE Mite group 1 protein; methyltrophic yeast; Escherichia coli; allergy;

XX recombinant mite group 1 protein; allergic response; antiallergic;

KW infectious disease; allergic disease.

XX Dermatophagoides pteronyssinus.

XX WO200129078-A2.

XX 26-APR-2001.

XX 12-OCT-2000; 2000WO-US028204.

XX 15-OCT-1999; 99US-0159841P.

XX (HESK-) HESKA CORP.

XX Best EA, McDermott MJ;

XX WPI; 2001-308475/32.

DR

DR N-PSDB; AAB22385.

XX Producing recombinant mite Group 1 protein for treating allergies,

XX involves culturing a methyltrophic yeast microorganism or Escherichia

PT coli transformed with nucleic acid molecule, and recovering the protein.

XX Claim 12; Page 145; 154pp; English.

XX The present invention describes a method for the production of a

CC recombinant mite Group 1 protein (I). The method comprises culturing a

CC methyltrophic yeast microorganism transformed with a nucleic acid

CC molecule (II) encoding (I), and recovering (I), or culturing Escherichia

CC coli transformed with (II) under conditions in which (I) forms an

CC inclusion body in E. coli, isolating the inclusion body, and recovering

CC comprising: (a) contacting (I) with a putative IgE-containing substance

CC to form a complex between (I) and IgE; and (b) determining the presence of

CC of IgE reactive with (I) by detecting the complex, where the presence of

CC reactive IgE is indicative of mite allergy in the animal. (I) is useful

CC for detecting mite allergy in an animal, or in a composition to reduce

CC allergic response to a mite Group 1 protein in a mite allergic animal.

CC (I) is also useful in a composition for treating or preventing allergic,

CC infectious or other diseases. AAB22326 to AAB22394 and AAB98326 to

CC AAB98349 represent sequences used in the exemplification of the present

CC invention

XX

SQ Sequence 222 AA;

Query Match 99.5%; Score 1200; DB 4; Length 222;

Best Local Similarity 99.5%; Pred. No. 1.3e-126;

Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQMRVTVPTRMQGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60

DB 1 TNACSGNAPAEIDLQMRVTVPTRMQGGCGSCWAFSGVAATESAYLAHRNQSILDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQSSYYRYVAREQSCRRPNAQRFGISNYCOIY 120

DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQSSYYRYVAREQSCRRPNAQRFGISNYCOIY 120

QY 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180

DB 121 PPNAKIREALAOHTSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYNSA 180

QY 181 QGVVDYWIVRNSWDNTNWDNGNGYGYFAANIDLMWIEEYPVWIL 222

DB 181 QGVVDYWIVRNSWDNTNWDNGNGYGYFAANIDLMWIEEYPVWIL 222

RESULT 5

AAU07746

ID AAU07746 standard; protein; 222 AA.

XX AAU07746;

XX 04-DEC-2001 (first entry)

XX House dust mite allergenic protein Der p I variant b.

XX House dust mite; allergenic protein; Der p I; Der p II; Der f I;

KW Der f II; antiallergic; immunostimulant; house dust mite allergy;

XX T-cell epitope; polymorphic variant.

XX Dermatophagoides pteronyssinus.

XX Key Location/Qualifiers

FT Misc-difference 50

FT /note= "Wild-type His substituted by Tyr"

FT Misc-difference 124

FT /note= "Wild-type Ala substituted by Val"

XX US6268491-B1.

XX

PD 31-JUL-2001.  
 XX 07-JUN-1995; 95US-00484296.  
 XX 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX (IMMU-) IMMUNOLOGIC PHARM CORP.  
 PA Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 XX Evans S, Shaked Z;  
 PI WPI; 2001-549074/61.  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX Disclosure; Fig 22; 158pp; English.  
 XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic protein from Dermatophagoides from  
 CC which the T-cell epitope containing peptides are derived, a polymorphic  
 CC variant of Der p I. Note: The present sequence is not shown in the  
 CC specification but is derived from the Der p I sequence shown in figure 22  
 XX Sequence 222 AA;  
 SQ  
 Query Match 99.2%; Score 1196; DB 4; Length 222;  
 Best Local Similarity 99.1%; Pred. No. 3.6e-126;  
 Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 QY 61 ELVDCASQHCCHGDTIPRGIEYIQHGVVQESYRVYVAREQSCRRPNAQRFGISNYCOIY 120  
 DB 61 ELVDCASQHCCHGDTIPRGIEYIQHGVVQESYRVYVAREQSCRRPNAQRFGISNYCOIY 120  
 QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIGYSNA 180  
 DB 121 PPNVKNKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIGYSNA 180  
 QY 181 QGVDDYIVRNSWDTNMGDNGYGFYFAANIDLMWIEEYPPVVIL 222  
 DB 181 QGVDDYIVRNSWDTNMGDNGYGFYFAANIDLMWIEEYPPVVIL 222  
 RESULT 6  
 ABG67023  
 ID ABG67023 standard; protein; 222 AA.  
 XX AC ABG67023;  
 XX AC ABG67023;  
 DT 24-SEP-2002 (first entry)  
 XX House dust mite allergen Der p 1.  
 DE

XX Immunoglobulin E; IGE; allergen; allergy; hay fever; rhinoconjunctivitis;  
 KW rhinitis; asthma; systemic anaphylaxis; vaccine; antiallergic;  
 KW B cell epitope.  
 XX Dermatophagoides pteronyssinus.  
 OS WO200240676-A2.  
 PN 23-MAY-2002.  
 XX 16-NOV-2001; 2001WO-DK000764.  
 PF 16-NOV-2000; 2000DK-00001718.  
 XX 16-NOV-2000; 2000US-0249361P.  
 PR 14-JUN-2001; 2001US-0298170P.  
 XX (ALKA-) ALK-ABELLO AS.  
 PA Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;  
 PI WPI; 2002-508328/54.  
 XX N-PSDB; ABK95639.  
 DR New recombinant mutant allergen, useful for preventing and/or treating  
 XX allergy, comprises multiple mutations and reduced immunoglobulin E  
 PT binding affinity.  
 PT Example 7; Page 103-104; 210pp; English.  
 XX The invention relates to a recombinant allergen (I) which is a mutant of  
 CC a naturally occurring allergen, where the mutant allergen has at least  
 CC four primary mutations, which each reduce the specific immunoglobulin E  
 CC (IGE) binding capability of the mutated allergen as compared to the IGE  
 CC binding capability of the naturally occurring allergen, where each  
 CC primary mutation is a substitution of one surface-exposed amino acid  
 CC residue with another residue, which does not occur in the same position  
 CC in the amino acid sequence of any known homologous protein within the  
 CC taxonomic species from which the naturally occurring allergen originates,  
 CC and each primary mutation is spaced from each other primary mutation by  
 CC at least 15 Angstrom, and the primary mutations are placed in such a  
 CC manner that at least one circular surface region with a area of 800  
 CC Angstrom<sup>2</sup> comprises no mutation. Also included are a composition  
 CC comprising two or more of the recombinant allergens, where the variant  
 CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, or a  
 CC sequence which hybridises to it under stringent conditions, where the  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IGE containing sample of the subject is mixed with  
 CC the recombinant allergen and assessed for the level of reactivity between  
 CC the IGE in the sample and the recombinant allergen. The recombinant  
 CC allergen or compositions are useful for generating an immune response in  
 CC a subject, for vaccination or treatment of a subject or for the  
 CC treatment, prevention or alleviation of allergic reactions in a subject  
 CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 CC anaphylaxis. The present sequence represents a wild-type allergen of the  
 CC invention  
 XX Sequence 222 AA;  
 SQ  
 Query Match 99.2%; Score 1196; DB 5; Length 222;  
 Best Local Similarity 99.1%; Pred. No. 3.6e-126;  
 Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60

Db 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
QY 121 PPNAKIREALAOQTHSAIAIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNAKIREALAOQTHSAIAIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMLMIEEYPYVVIL 222  
Db 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMLMIEEYPYVVIL 222

RESULT 7  
ADK52140  
ID ADK52140 standard; protein; 222 AA.  
XX AC ADK52140;  
XX DT 06-MAY-2004 (first entry)  
XX DE Der p1 allergen.  
XX KW recombinant protein allergen; Antiallergic; Desensitization; antibody;  
XX KW allergy; house dust mite; allergen.  
XX OS Dermatophagoides pteronyssinus.  
XX PN WO2004005334-A2.  
XX PD 15-JAN-2004.  
XX PF 04-JUL-2003; 2003WO-FR002085.  
XX PR 05-JUL-2002; 2002FR-00008485.  
XX PA (STAL-) STALLERGENES SA.  
XX PA (SETB ) SOC NAT EXPL IND TABACS & ALLUMETTES.  
XX PI Gomord V, Lienard D, Van Ree R, Van Oort E, Dorlhac De Borne F;  
PI Didier Laurent A, Faye L;  
DR WPI: 2004-083498/08.  
DR N-PSDB; ADK52139.  
XX Recombinant production of acarid protein allergen, useful for diagnosis  
PT and treatment of allergy to house dust mites, comprises growing  
PT transformed eukaryotes, particularly plants.  
XX Claim 4; SEQ ID NO 2; 55pp; French.  
XX The present invention relates to a method for production of a recombinant  
CC protein allergen from an acarid of the genera Dermatophagoides or  
CC Euroglyphus. The allergens and also antibodies raised against them, are  
CC useful for diagnosis and treatment of allergies to house dust mites. When  
CC expressed in plants, allergens are synthesized and matured to  
CC biologically active form, with essentially the same pattern of  
CC glycosylation as the native protein. Recombinant expression provides a  
CC pure protein; contrast complex mixtures of allergens currently used. The  
CC present sequence represents Der p1 allergen.  
XX Sequence 222 AA;

Query Match 99.2%; Score 1196; DB 8; Length 222;  
Best Local Similarity 99.1%; Pred. No. 3.6e-126;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
QY 121 PPNAKIREALAOQTHSAIAIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNAKIREALAOQTHSAIAIIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMLMIEEYPYVVIL 222  
Db 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMLMIEEYPYVVIL 222

RESULT 8  
ADR87225  
ID ADR87225 standard; protein; 222 AA.  
XX AC ADR87225;  
XX DT 16-DEC-2004 (first entry)  
XX DE Dust mite allergen Der p 1 SEQ ID NO:14.  
XX KW dust mite; allergen; Der p 1; T-cell response; IgE; immunoglobulin E;  
XX KW immune response; antiallergic; gene therapy; vaccine.  
XX OS Dermatophagoides sp.  
XX PN WO2004081028-A2.  
XX PD 23-SEP-2004.  
XX PF 15-MAR-2004; 2004WO-IB001300.  
XX PR 14-MAR-2003; 2003US-0455004P.  
XX PR 12-MAR-2004; 2004US-00799514.  
XX PA (UYLA-) UNIV LAUSANNE.  
XX PI Spertini F;  
XX WPI: 2004-668931/65.  
XX New compositions including contiguous overlapping peptide fragments that  
PT form an entire amino acid sequence of an allergen (e.g. bee venom or  
PT birch pollen allergen), useful for preventing or treating IgE-mediated  
PT allergies.  
XX Claim 4; SEQ ID NO 14; 82pp; English.  
XX The invention relates to novel compositions including contiguous  
CC overlapping peptide fragments which together form an entire amino acid  
CC sequence of an allergen, where the fragments are capable of inducing a T-  
CC cell response in patients who are hypersensitive to the allergen. The  
CC contiguous overlapping peptide fragments further result in lower levels  
CC of IgE stimulation activity. The lower levels of IgE stimulation activity  
CC are zero or weak. The contiguous overlapping peptide fragments further  
CC result in a decrease in T-cell response upon subsequent exposure to the  
CC allergen, thus, modulating an immune response in the patients, who are  
CC hypersensitive to the allergen. A composition of the invention has  
CC antiallergic activity, and may have a use in gene therapy, and as a  
CC vaccine. The composition and methods are useful for preventing or  
CC treating IgE-mediated allergies. The present sequence represents an  
CC allergen of the invention, dust mite Der p 1.  
XX Sequence 222 AA;

Query Match 99.2%; Score 1196; DB 8; Length 222;  
Best Local Similarity 99.1%; Pred. No. 3.6e-126;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TNACSSINGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60

DB 1 TNACSLNGNAPAEIDLQRMRTVTPIRMQGGSCSWAFSGVAATESAYLAHRNQLDLAEQ 60  
QY 61 ELVDCAHQCHGCHGDTTPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 61 ELVDCAHQCHGCHGDTTPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHVDGRTIIQRDNGYOPNVHNAVINGYSNA 180  
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHVDGRTIIQRDNGYOPNVHNAVINGYSNA 180  
QY 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMMEIEEYPYVIL 222  
DB 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMMEIEEYPYVIL 222

RESULT 9  
AAU07747  
ID AAU07747 standard; protein; 222 AA.  
AC AAU07747;  
XX  
XX  
DT 04-DEC-2001 (first entry)  
DE House dust mite allergenic protein Der p I variant c.  
KW House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
KW Der f II; anti-allergenic; immunostimulant; house dust mite allergy;  
KW T-cell epitope; polymorphic variant.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 50 /note= "Wild-type His substituted by Tyr"  
FT Misc-difference 81 /note= "Wild-type Glu substituted by Lys"  
FT Misc-difference 124 /note= "Wild-type Ala substituted by Val"  
FT Misc-difference 215 /note= "Wild-type Glu substituted by Gln"  
XX  
XX US6268491-B1.  
PN  
XX  
PD 31-JUL-2001.  
XX  
XX 07-JUN-1995; 95US-00484296.  
XX  
XX 16-OCT-1991; 91US-00777859.  
PR 08-MAY-1992; 92US-00881396.  
PR 14-APR-1993; 93WO-US003471.  
PR 14-APR-1994; 94US-00227772.  
PR 19-MAY-1995; 95US-00445307.  
XX  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX  
XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
PI Evans S, Shaked Z;  
XX  
XX WPI; 2001-549074/61.  
DR  
XX Peptides comprising T cell groups of the major allergens from  
PT Dermatophagoides (house dust mites), useful for treating house dust mite  
PT allergy in humans, and for diagnosing sensitivity to house dust mite  
PT protein allergens.  
XX  
XX Disclosure; Fig 22; 158pp; English.  
PS  
XX The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions, each  
CC region comprising at least one T cell group of a protein allergen of the

CC genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced  
CC therapeutic properties as the naturally-occurring allergen, but have  
CC reduced side effects, and increased solubility and stability. The present  
CC sequence represents an allergenic protein from Dermatophagoides from  
CC which the T-cell epitope containing peptides are derived, a polymorphic  
CC variant of Der p I. Note: The present sequence is not shown in the  
CC specification but is derived from the Der p I sequence shown in figure 22  
XX  
XX Sequence 222 AA;  
QY  
XX  
Query Match 98.6%; Score 1189; DB 4; Length 222;  
Best Local Similarity 98.2%; Pred. No. 2.2e-125;  
Matches 218; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TNACSLNGNAPAEIDLQRMRTVTPIRMQGGSCSWAFSGVAATESAYLAHRNQLDLAEQ 60  
DB 1 TNACSLNGNAPAEIDLQRMRTVTPIRMQGGSCSWAFSGVAATESAYLAHRNQLDLAEQ 60  
QY 61 ELVDCAHQCHGCHGDTTPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 61 ELVDCAHQCHGCHGDTTPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120  
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHVDGRTIIQRDNGYOPNVHNAVINGYSNA 180  
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHVDGRTIIQRDNGYOPNVHNAVINGYSNA 180  
QY 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMMEIEEYPYVIL 222  
DB 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMMEIEEYPYVIL 222

RESULT 10  
AAR52742  
ID AAR52742 standard; protein; 222 AA.  
XX  
XX AAR52742;  
AC  
XX  
DT 25-MAR-2003 (revised)  
DT 17-OCT-1994 (first entry)  
XX  
XX Protein allergen of Der pI.  
XX  
XX Der pI; House Dust Mite Allergen.  
XX  
XX Dermatophagoides pteronyssinus.  
OS  
XX Key Location/Qualifiers  
FT Misc-difference 50 /label= His or Tyr  
FT Misc-difference 81 /label= Glu or Lys  
FT Misc-difference 124 /label= Ala or Val  
FT Misc-difference 136 /label= Ser or Thr  
FT Misc-difference 215 /label= Glu or Gln  
XX  
XX WO9405790-A1.  
PN  
XX  
XX 17-MAR-1994.  
PD  
XX  
XX 10-SEP-1993; 93WO-US008518.  
PF  
XX  
XX 10-SEP-1992; 92US-00945288.  
PR  
XX  
XX (IMMU-) IMMULOGIC PHARM CORP.  
PA  
XX  
XX Thomas WR, Chua K;  
PI

XX WPI; 1994-101195/12.  
 XX  
 PT New protein allergens of house dust mite - used for diagnosing and  
 PT treating sensitivity in an individual to house dust mite allergens.  
 XX  
 XX Disclosure; Fig 18; 98pp; English.  
 XX  
 CC AAR52742 is a mature Der pI mite allergen. The mature protein can be used  
 CC to detect sensitivity in an individual to house dust mite and to reduce  
 CC the sensitivity of the individual. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX  
 SQ Sequence 222 AA;  
 Query Match 97.6%; Score 1177; DB 2; Length 222;  
 Best Local Similarity 97.7%; Pred. No. 5.1e-124;  
 Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLRQMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 1 TNACSSINGNAPAEIDLRQMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVWQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIXYIQHNGVWQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
 DB 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
 QY 181 QGVYDIWVRNSWDTNWDNGYGYFAANIDLMIEEYPYVIL 222  
 DB 181 QGVYDIWVRNSWDTNWDNGYGYFAANIDLMIEEYPYVIL 222  
 RESULT 11  
 ABG67024  
 ID ABG67024 standard; protein; 222 AA.  
 XX  
 AC ABG67024;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 DE House dust mite allergen Der p 1 mutant #1.  
 DE  
 KW Immunoglobulin E; IgE; allergen; allergy; mite; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 OS Synthetic.  
 XX  
 PN WO200240676-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 16-NOV-2001; 2001WO-DK000764.  
 XX  
 PR 16-NOV-2000; 2000DK-00001718.  
 PR 16-NOV-2000; 2000US-0249361P.  
 PR 14-JUN-2001; 2001US-0298170P.  
 XX  
 PA (ALKA-) ALK-ABELLO AS.  
 XX  
 PI Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;  
 XX  
 XX WPI; 2002-508328/54.  
 DR N-PSDB; ABK95640.  
 XX  
 XX New recombinant mutant allergen, useful for preventing and/or treating  
 PT allergy, comprises multiple mutations and reduced immunoglobulin E  
 PT binding affinity.

XX  
 PS  
 XX  
 CC The invention relates to a recombinant allergen (I) which is a mutant of  
 CC a naturally occurring allergen, where the mutant allergen has at least  
 CC four primary mutations, which each reduce the specific immunoglobulin E  
 CC (IgE) binding capability of the mutated allergen as compared to the IgE  
 CC binding capability of the naturally occurring allergen, where each  
 CC primary mutation is a substitution of one surface-exposed amino acid  
 CC residue with another residue, which does not occur in the same position  
 CC in the amino acid sequence of any known homologous protein within the  
 CC taxonomic species from which the naturally occurring allergen originates,  
 CC and each primary mutation is spaced from each other primary mutation by  
 CC at least 15 Angstrom , and the primary mutations are placed in such a  
 CC manner that at least one circular surface region with a area of 800  
 CC Angstrom ^2 comprises no mutation. Also included are a composition  
 CC comprising two or more of the recombinant allergens, where the variant  
 CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, or a  
 CC sequence which hybridises to it under stringent conditions, where the  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IgE containing sample of the subject is mixed with  
 CC the recombinant allergen and assessed for the level of reactivity between  
 CC the IgE in the sample and the recombinant allergen. The recombinant  
 CC allergen or compositions are useful for generating an immune response in  
 CC a subject, for vaccination or treatment of a subject or for the  
 CC treatment, prevention or alleviation of allergic reactions in a subject  
 CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 CC anaphylaxis. The present sequence represents a recombinant allergen of  
 CC the invention. Note: The present sequence was not shown in the  
 CC specification but was created by the indexer using information in the  
 CC specification and the corresponding wild-type sequence  
 XX  
 SQ Sequence 222 AA;  
 Query Match 96.4%; Score 1162; DB 5; Length 222;  
 Best Local Similarity 96.8%; Pred. No. 2.5e-122;  
 Matches 215; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TNACSSINGNAPAEIDLRQMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 1 TNACSSINGNAPAEIDLRQMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVWQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPQIEYIQHNGVWQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
 DB 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYSNA 180  
 QY 181 QGVYDIWVRNSWDTNWDNGYGYFAANIDLMIEEYPYVIL 222  
 DB 181 QGVYDIWVRNSWDTNWDNGYGYFAANIDLMIEEYPYVIL 222  
 RESULT 12  
 ABG67030  
 ID ABG67030 standard; protein; 222 AA.  
 XX  
 AC ABG67030;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 DE House dust mite allergen Der p 1 mutant #7.

XX Immunoglobulin E; IgE; allergen; allergy; mutein; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 OS Synthetic.  
 OS WO200240676-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 16-NOV-2001; 2001WO-DK000764.  
 XX  
 PR 16-NOV-2000; 2000DK-00001718.  
 PR 16-NOV-2000; 2000US-0249361P.  
 PR 14-JUN-2001; 2001US-0298170P.  
 XX  
 XX (ALKA-) ALK-ABELLO AS.  
 XX  
 XX Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;  
 XX  
 XX WPI; 2002-508328/54.  
 XX N-PSDB; ABK95646.  
 XX  
 XX New recombinant mutant allergen, useful for preventing and/or treating  
 PT allergy, comprises multiple mutations and reduced immunoglobulin E  
 PT binding affinity.  
 XX  
 XX Example 7; Page: 210pp; English.  
 XX  
 XX The invention relates to a recombinant allergen (I) which is a mutant of  
 CC a naturally occurring allergen, where the mutant allergen has at least  
 CC four primary mutations, which each reduce the specific immunoglobulin E  
 CC (IgE) binding capability of the mutated allergen as compared to the IgE  
 CC binding capability of the naturally occurring allergen, where each  
 CC primary mutation is a substitution of one surface-exposed amino acid  
 CC residue with another residue, which does not occur in the same position  
 CC in the amino acid sequence of any known homologous protein within the  
 CC taxonomic species from which the naturally occurring allergen originates,  
 CC and each primary mutation is spaced from each other primary mutation by  
 CC at least 15 Angstrom , and the primary mutations are placed in such a  
 CC manner that at least one circular surface region with a area of 800  
 CC Angstrom ^2 comprises no mutation. Also included are a composition  
 CC comprising two or more of the recombinant allergens, where the variant  
 CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, where the  
 CC sequence which hybridises to it under stringent conditions, where the  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IgE containing sample of the subject is mixed with  
 CC the recombinant allergen and assessed for the level of reactivity between  
 CC the IgE in the sample and the recombinant allergen. The recombinant  
 CC allergen or compositions are useful for generating an immune response in  
 CC a subject, for vaccination or treatment of a subject or for the  
 CC treatment, prevention or alleviation of allergic reactions in a subject  
 CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 CC anaphylaxis. The present sequence represents a recombinant allergen of  
 CC the invention. Note: The present sequence was not shown in the  
 CC specification but was created by the indexer using information in the  
 CC specification and the corresponding wild-type sequence

Sequence 222 AA;

Query Match 96.4%; Score 1162; DB 5; Length 222;  
 Best Local Similarity 96.4%; Pred. No. 2.5e-122;

Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TNACSGNGAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 60  
 DB 1 TNACSGNGAPAEIDLQRMQVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQLDLAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGLSNYQIY 120  
 DB 61 ELVDCANQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGLSNYQIY 120  
 QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIORNGYOPFNHAVNIVYSNA 180  
 DB 121 PPNVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGETIIQEDNGYQTNHAVNIVYSNA 180  
 QY 181 QGVGYWIVRNSWDTNWGDNNGYGYFAANIDLMIMEEYPYVUIL 222  
 DB 181 QGVGYWIVRNSWDTNWGDNNGYGYFAANIDLMIMEEYPYVUIL 222  
 RESULT 13  
 ABG67027  
 ID ABG67027 standard; protein; 222 AA.  
 XX  
 XX AC ABG67027;  
 XX  
 XX DT 24-SEP-2002 (first entry)  
 XX  
 XX DE House dust mite allergen Der p 1 mutant #4.  
 XX  
 KW Immunoglobulin E; IgE; allergen; allergy; mutein; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.  
 XX  
 XX Dermatophagoides pteronyssinus.  
 OS Synthetic.  
 XX WO200240676-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 16-NOV-2001; 2001WO-DK000764.  
 XX  
 PR 16-NOV-2000; 2000DK-00001718.  
 PR 16-NOV-2000; 2000US-0249361P.  
 PR 14-JUN-2001; 2001US-0298170P.  
 XX  
 XX (ALKA-) ALK-ABELLO AS.  
 XX  
 XX Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;  
 XX  
 XX WPI; 2002-508328/54.  
 XX N-PSDB; ABK95643.  
 XX  
 XX New recombinant mutant allergen, useful for preventing and/or treating  
 PT allergy, comprises multiple mutations and reduced immunoglobulin E  
 PT binding affinity.  
 XX  
 XX Example 7; Page: 210pp; English.  
 XX  
 XX The invention relates to a recombinant allergen (I) which is a mutant of  
 CC a naturally occurring allergen, where the mutant allergen has at least  
 CC four primary mutations, which each reduce the specific immunoglobulin E  
 CC (IgE) binding capability of the mutated allergen as compared to the IgE  
 CC binding capability of the naturally occurring allergen, where each  
 CC primary mutation is a substitution of one surface-exposed amino acid  
 CC residue with another residue, which does not occur in the same position  
 CC in the amino acid sequence of any known homologous protein within the  
 CC taxonomic species from which the naturally occurring allergen originates,  
 CC and each primary mutation is spaced from each other primary mutation by  
 CC at least 15 Angstrom , and the primary mutations are placed in such a  
 CC manner that at least one circular surface region with a area of 800  
 CC Angstrom ^2 comprises no mutation. Also included are a composition  
 CC comprising two or more of the recombinant allergens, where the variant  
 CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, where the  
 CC sequence which hybridises to it under stringent conditions, where the  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IgE containing sample of the subject is mixed with  
 CC the recombinant allergen and assessed for the level of reactivity between  
 CC the IgE in the sample and the recombinant allergen. The recombinant  
 CC allergen or compositions are useful for generating an immune response in  
 CC a subject, for vaccination or treatment of a subject or for the  
 CC treatment, prevention or alleviation of allergic reactions in a subject  
 CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 CC anaphylaxis. The present sequence represents a recombinant allergen of  
 CC the invention. Note: The present sequence was not shown in the  
 CC specification but was created by the indexer using information in the  
 CC specification and the corresponding wild-type sequence



CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, where the  
 CC sequence which hybridises to it under stringent conditions, or a  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IgE containing sample of the subject is mixed with  
 CC the recombinant allergen and assessed for the level of reactivity between  
 CC the IgE in the sample and the recombinant allergen. The recombinant  
 CC allergen or compositions are useful for generating an immune response in  
 CC a subject, for vaccination or treatment of a subject or for the  
 CC treatment, prevention or alleviation of allergic reactions in a subject  
 CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 CC anaphylaxis. The present sequence represents a recombinant allergen of  
 CC the invention. Note: The present sequence was not shown in the  
 CC specification but was created by the indexer using information in the  
 CC specification and the corresponding wild-type sequence

XX Sequence 222 AA;

Query Match 96.3%; Score 1161; DB 5; Length 222;

Best Local Similarity 96.4%; Pred. No. 3.2e-122;

Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLROMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 1 TNACSSINGNAPAEIDLROMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 QY 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMIEEYPYVIL 222  
 DB 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMIEEYPYVIL 222

RESULT 14

ABG67029  
 ID ABG67029 standard; protein; 222 AA.

XX AC ABG67029;

XX DT 24-SEP-2002 (first entry)

XX DE House dust mite allergen Der p 1 mutant #6.

XX KW Immunoglobulin E; IgE; allergen; allergy; mutein; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.

XX OS Dermotophagoides pteronyssinus.  
 OS Synthetic.

XX PN W0200240676-A2.

XX PD 23-MAY-2002.

XX PR 16-NOV-2001; 2001WO-DK000764.

XX PR 16-NOV-2000; 2000DK-00001718.

XX PR 16-NOV-2000; 2000US-0249361P.

XX PR 14-JUN-2001; 2001US-0298170P.

(ALKA-) ALK-ABELLO AS.

Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;

WPI; 2002-508328/54.

N-PSDB; ABK95645.

New recombinant mutant allergen, useful for preventing and/or treating  
 allergy, comprises multiple mutations and reduced immunoglobulin E  
 binding affinity.

Example 7; Page: 210pp; English.

The invention relates to a recombinant allergen (I) which is a mutant of  
 a naturally occurring allergen, where the mutant allergen has at least  
 four primary mutations, which each reduce the specific immunoglobulin E  
 (IgE) binding capability of the mutated allergen as compared to the IgE  
 binding capability of the naturally occurring allergen, where each  
 primary mutation is a substitution of one surface-exposed amino acid  
 residue with another residue, which does not occur in the same position  
 in the amino acid sequence of any known homologous protein within the  
 taxonomic species from which the naturally occurring allergen originates,  
 and each primary mutation is spaced from each other primary mutation by  
 at least 15 Angstrom, and the primary mutations are placed in such a  
 manner that at least one circular surface region with a area of 800  
 Angstrom<sup>2</sup> comprises no mutation. Also included are a composition  
 comprising two or more of the recombinant allergens, where the variant  
 allergen is defined by having at least one primary mutation, which is  
 absent in at least one of the other variants, and for each variant no  
 secondary mutation is present within a radius of 15 Angstrom from each  
 absent primary mutation; a DNA sequence encoding the recombinant allergen  
 or its derivative, partial sequence or degenerated sequence, or a  
 sequence which hybridises to it under stringent conditions, where the  
 derivative, partial sequence, degenerated sequence or hybridising  
 sequence encodes a peptide having at least one B cell epitope; an  
 expression vector comprising the DNA and a host cell comprising the  
 vector. The recombinant allergen is useful as a pharmaceutical, for  
 preparing a pharmaceutical for preventing and/or treating allergy, or in  
 a diagnostic assay for assessing relevance, safety or outcome of therapy  
 of a subject, where an IgE containing sample of the subject is mixed with  
 the recombinant allergen and assessed for the level of reactivity between  
 the IgE in the sample and the recombinant allergen. The recombinant  
 allergen or compositions are useful for generating an immune response in  
 a subject, for vaccination or treatment of a subject or for the  
 treatment, prevention or alleviation of allergic reactions in a subject  
 e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
 anaphylaxis. The present sequence represents a recombinant allergen of  
 the invention. Note: The present sequence was not shown in the  
 specification but was created by the indexer using information in the  
 specification and the corresponding wild-type sequence

XX Sequence 222 AA;

Query Match 96.3%; Score 1161; DB 5; Length 222;

Best Local Similarity 96.4%; Pred. No. 3.2e-122;

Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLROMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 1 TNACSSINGNAPAEIDLROMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 QY 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMIEEYPYVIL 222  
 DB 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMIEEYPYVIL 222

## RESULT 15

ABG67026  
ID ABG67026 standard; protein; 222 AA.  
XX AC ABG67026;  
XX DT 24-SEP-2002 (first entry)  
XX DE House dust mite allergen Der p 1 mutant #3.  
XX KW Immunoglobulin E; IGE; allergen; allergy; mite; hay fever;  
XX KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
XX KW vaccine; antiallergic; B cell epitope.  
XX OS Dermatophagoides pteronyssinus.  
XX OS Synthetic.  
XX PN WC200240676-A2.  
XX PD 23-MAY-2002.  
XX PF 16-NOV-2001; 2001WO-DK000764.  
XX PR 16-NOV-2000; 2000DK-00001718.  
XX PR 16-NOV-2000; 2000US-0249361P.  
XX PR 14-JUN-2001; 2001US-0298170P.  
XX PA (ALKA-) ALK-ABELLO AS.  
XX PI Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;  
XX WPI; 2002-508328/54.  
XX DR N-PSDB; ABK95642.

XX New recombinant mutant allergen, useful for preventing and/or treating  
PT allergy, comprises multiple mutations and reduced immunoglobulin E  
PT binding affinity.

XX Example 7; Page; 210pp; English.

XX The invention relates to a recombinant allergen (I) which is a mutant of  
CC a naturally occurring allergen, where the mutant allergen has at least  
CC four primary mutations, which each reduce the specific immunoglobulin E  
CC (IgE) binding capability of the mutated allergen as compared to the IgE  
CC binding capability of the naturally occurring allergen, where each  
CC primary mutation is a substitution of one surface-exposed amino acid  
CC residue with another residue, which does not occur in the same position  
CC in the amino acid sequence of any known homologous protein within the  
CC taxonomic species from which the naturally occurring allergen originates,  
CC and each primary mutation is spaced from each other primary mutation by  
CC at least 15 Angstrom , and the primary mutations are placed in such a  
CC manner that at least one circular surface region with a area of 800  
CC Angstrom ^2 comprises no mutation. Also included are a composition  
CC comprising two or more of the recombinant allergens, where the variant  
CC allergen is defined by having at least one primary mutation, which is  
CC absent in at least one of the other variants, and for each variant no  
CC secondary mutation is present within a radius of 15 Angstrom from each  
CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
CC or its derivative, partial sequence or degenerated sequence, or a  
CC sequence which hybridises to it under stringent conditions, where the  
CC derivative, partial sequence, degenerated sequence or hybridising  
CC sequence encodes a peptide having at least one B cell epitope; an  
CC expression vector comprising the DNA and a host cell comprising the  
CC vector. The recombinant allergen is useful as a pharmaceutical, for  
CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
CC of a subject, where an IGE containing sample of the subject is mixed with  
CC the recombinant allergen and assessed for the level of reactivity between  
CC the IGE in the sample and the recombinant allergen. The recombinant  
CC allergen or compositions are useful for generating an immune response in  
CC a subject, for vaccination or treatment of a subject or for the

CC treatment, prevention or alleviation of allergic reactions in a subject  
CC e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or systemic  
CC anaphylaxis. The present sequence represents a recombinant allergen of  
CC the invention. Note: The present sequence was not shown in the  
CC specification but was created by the indexer using information in the  
CC specification and the corresponding wild-type sequence

XX SQ Sequence 222 AA;

Query Match 96.2%; Score 1160; DB 5; Length 222;  
Best Local Similarity 96.4%; Pred. NO. 4.2e-122;  
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSTNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

Db 1 TNACSTNGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

QY 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRONGYQPNTHAVNIVGYSNA 180

Db 121 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRONGYQPNTHAVNIVGYSNA 180

QY 181 QGVVDYWIVERNWDNTNWDNGYGFYFAANIDLMMEIEYPYVIL 222

Db 181 QGVVDYWIVERNWDNTNWDNGYGFYFAANIDLMMEIEYPYVIL 222

Search completed: May 19, 2005, 17:39:01

Job time : 163 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:36:23 ; Search time 135 Seconds  
(without alignments)  
550.079 Million cell updates/sec

Title: US-09-867-159A-2  
Perfect score: 1206  
Sequence: 1 TNACSSINGNAPAEIDLQRM.....YFAANIDLMIEBYPVVIL 222

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 893949

Minimum DB seq length: 0  
Maximum DB seq length: 222

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206	100.0	222	10	US-09-867-159A-2
2	1196	99.2	222	14	US-10-001-245-88
3	1162	96.4	222	14	US-10-001-245-26
4	1161	96.3	222	14	US-10-001-245-18
5	1161	96.3	222	14	US-10-001-245-20
6	1161	96.3	222	14	US-10-001-245-24
7	1160	96.2	218	17	US-10-892-543-5
8	1159	96.1	222	14	US-10-001-245-22
9	1157	95.9	222	14	US-10-001-245-14
10	1157	95.9	222	14	US-10-001-245-16
11	1143	94.8	222	14	US-10-001-245-30
12	1142	94.7	222	14	US-10-001-245-28
13	1133	93.9	222	14	US-10-001-245-34

14	1132	93.9	222	14	US-10-001-245-32	Sequence 32, Appl
15	982.5	81.5	211	10	US-09-847-208-95	Sequence 95, Appl
16	982.5	81.5	211	14	US-10-001-245-184	Sequence 184, App
17	944.5	78.3	210	14	US-10-001-245-185	Sequence 185, App
18	303	25.1	217	15	US-10-360-101-215	Sequence 215, App
19	297.5	24.7	218	11	US-09-972-211-90	Sequence 90, Appl
20	297.5	24.7	218	15	US-10-096-625-90	Sequence 90, Appl
21	294.5	24.4	220	11	US-09-972-211-89	Sequence 89, Appl
22	294.5	24.4	220	15	US-10-096-625-89	Sequence 89, Appl
23	282.5	23.4	210	16	US-10-466-110-7	Sequence 7, Appli
24	280.5	23.3	222	17	US-10-635-398-98	Sequence 98, Appl
25	278.5	23.1	217	14	US-10-273-577-5	Sequence 5, Appli
26	274	22.7	201	15	US-10-425-114-42955	Sequence 42955, A
27	271.5	22.5	217	14	US-10-273-577-2	Sequence 2, Appli
28	268.5	22.3	217	14	US-10-273-577-4	Sequence 4, Appli
29	261.5	21.7	217	14	US-10-273-577-3	Sequence 3, Appli
30	254	21.1	220	9	US-09-462-846-3	Sequence 3, Appli
31	254	21.1	220	16	US-10-773-387-3	Sequence 3, Appli
32	254	21.1	220	16	US-10-773-914-3	Sequence 3, Appli
33	251.5	20.9	200	14	US-10-219-220-150	Sequence 150, App
34	247	20.5	212	16	US-10-872-198-22	Sequence 22, Appl
35	247	20.5	212	17	US-10-872-197A-22	Sequence 22, Appl
36	235	19.5	205	15	US-10-425-114-53588	Sequence 53588, A
37	211	17.5	198	14	US-10-219-220-151	Sequence 151, App
38	208	17.2	129	14	US-10-219-220-159	Sequence 159, App
39	206	17.1	193	15	US-10-425-114-46437	Sequence 46437, A
40	205.5	17.0	195	16	US-10-767-701-43624	Sequence 43624, A
41	204	16.9	169	16	US-10-767-701-36832	Sequence 36832, A
42	201	16.7	147	14	US-10-219-220-173	Sequence 173, App
43	199.5	16.5	163	16	US-10-767-701-45098	Sequence 45098, A
44	197.5	16.4	160	14	US-10-219-220-155	Sequence 155, App
45	196	16.3	185	15	US-10-425-114-57524	Sequence 57524, A

ALIGNMENTS

RESULT 1

US-09-867-159A-2  
; Sequence 2, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antialis and at least one anti-histamine compound  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteromyssinus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)...(222)  
; OTHER INFORMATION: Peptide sequence from cystine protease.  
US-09-867-159A-2

Query Match	100.0%	Score 1206;	DB 10;	Length 222;
Best Local Similarity	100.0%;	Pred. No. 1.3e-118;		
Matches	222;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Oy	1	TNACSSINGNAPAEIDLQRMVTVPTRMQGGCGSWAFSGVAATPSAYLAHRNQSILDLAEQ	60	
Db	1	TNACSSINGNAPAEIDLQRMVTVPTRMQGGCGSWAFSGVAATPSAYLAHRNQSILDLAEQ	60	
Oy	61	ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNQRFGISNYCQIY	120	

Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
QY 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222  
Db 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222

## RESULT 2

US-10-001-245-88  
; Sequence 88, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 88  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-88

Query Match 99.2%; Score 1196; DB 14; Length 222;  
Best Local Similarity 99.1%; Pred. No. 1.4e-117; Indels 0; Gaps 0;  
Matches 220; Conservative 1; Mismatches 1;  
QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
QY 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222  
Db 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222

## RESULT 3

US-10-001-245-26  
; Sequence 26, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-26

Query Match 96.4%; Score 1162; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 5.5e-114;  
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
QY 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222  
Db 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222

## RESULT 4

US-10-001-245-18  
; Sequence 18, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-18

Query Match 96.3%; Score 1161; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 7e-114; Indels 0; Gaps 0;  
Matches 214; Conservative 3; Mismatches 5;  
QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
Db 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGLSNYCOIY 120  
QY 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
Db 121 PPNANKIREALAOQTHSAIAVIGIKOLDAPFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVUIL 222

Db 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||||

## RESULT 5

US-10-001-245-20  
; Sequence 20, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-20

Query Match 96.3%; Score 1161; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 7e-114; Indels 0; Gaps 0;  
Matches 214; Conservative 3; Mismatches 5;  
Qy 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
|||  
Db 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
|||  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 120  
|||  
Db 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 120  
|||  
Qy 121 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 180  
|||  
Db 121 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 180  
|||  
Qy 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||  
Db 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||

## RESULT 6

US-10-001-245-24  
; Sequence 24, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus

US-10-001-245-24

Query Match 96.3%; Score 1161; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 7e-114; Indels 0; Gaps 0;  
Matches 214; Conservative 2; Mismatches 6;

Qy 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
|||  
Db 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
|||  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 120  
|||  
Db 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 120  
|||  
Qy 121 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 180  
|||  
Db 121 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 180  
|||  
Qy 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||  
Db 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||

## RESULT 7

US-10-892-543-5  
; Sequence 5, Application US/10892543  
; Publication No. US20050053615A1  
; GENERAL INFORMATION:  
; APPLICANT: Best, Elaine A.  
; APPLICANT: McDermott, Martin J.  
; TITLE OF INVENTION: VARIANTS OF MITE GROUP 1 ALLERGENS FOR THE TREATMENT OF HOUSE  
; FILE REFERENCE: AL-10  
; CURRENT APPLICATION NUMBER: US/10/892,543  
; CURRENT FILING DATE: 2004-07-15  
; PRIOR APPLICATION NUMBER: 60/487,812  
; PRIOR FILING DATE: 2003-07-16  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-892-543-5

Query Match 96.2%; Score 1160; DB 17; Length 218;  
Best Local Similarity 97.7%; Pred. No. 8.7e-114;  
Matches 217; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

Qy 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
|||  
Db 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 56  
|||  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 120  
|||  
Db 57 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFISNYCOIY 116  
|||  
Qy 121 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 180  
|||  
Db 117 PPNNKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHAVNIVGYNSA 176  
|||  
Qy 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
|||  
Db 177 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 218  
|||

## RESULT 8

US-10-001-245-22  
; Sequence 22, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik

```
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-22

Query Match          96.1%; Score 1159; DB 14; Length 222;
Best Local Similarity 96.4%; Pred. No. 1.1e-113;
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
DB 61 ELVDCANQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
DB 121 PPNVKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
QY 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
DB 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
; RESULT 9
US-10-001-245-14
; Sequence 14, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-14

Query Match          95.9%; Score 1157; DB 14; Length 222;
Best Local Similarity 96.4%; Pred. No. 1.8e-113;
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
DB 121 PPNVKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
QY 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
DB 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
; RESULT 10
US-10-001-245-16
; Sequence 16, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-16

Query Match          95.9%; Score 1157; DB 14; Length 222;
Best Local Similarity 96.4%; Pred. No. 1.8e-113;
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQCGSCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVQESYRYRYVAREQSCRRPNAQRFGISNYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
DB 121 PPNVKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHAVNIVGYSNA 180
QY 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
DB 181 QGVYWIYVNSWDTNWGNGYGYFAANIDLMIMEEYPYVIL 222
; RESULT 11
US-10-001-245-30
; Sequence 30, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
```

```
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-30

Query Match          94.8%; Score 1143; DB 14; Length 222;
Best Local Similarity 95.0%; Pred. No. 5.5e-112;
Matches 211; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 TNACSGNAPASIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
QY 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
QY 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
DB 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222
DB 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

RESULT 12
US-10-001-245-28
; Sequence 28, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-28

Query Match          94.7%; Score 1142; DB 14; Length 222;
Best Local Similarity 95.0%; Pred. No. 7e-112;
Matches 211; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 TNACSGNAPASIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
QY 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
QY 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
DB 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222
```

```
DB 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

RESULT 13
US-10-001-245-34
; Sequence 34, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-34

Query Match          93.9%; Score 1133; DB 14; Length 222;
Best Local Similarity 94.8%; Pred. No. 6.3e-111;
Matches 210; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 TNACSGNAPASIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
QY 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAQFGISNYCQIY 120
QY 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
DB 121 PPVANKIEEALQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
QY 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222
DB 181 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

RESULT 14
US-10-001-245-32
; Sequence 32, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:31:52 ; Search time 22 Seconds  
(without alignments)  
753.277 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSSINGNAPAEIDLQMR.....YFAANIDLMIEBPPYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 402398

Minimum DB seq length: 0

Maximum DB seq length: 222

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgm2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgm2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgm2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgm2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgm2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*

6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1177	97.6	222	1 US-07-945-288-11	Sequence 11, Appl
2	1177	97.6	222	1 US-08-462-831-11	Sequence 11, Appl
3	1177	97.6	222	1 US-08-461-809-11	Sequence 11, Appl
4	1177	97.6	222	1 US-08-461-441-11	Sequence 11, Appl
5	1177	97.6	222	5 PCT-US93-08518-11	Sequence 11, Appl
6	372.5	30.9	181	2 US-08-482-142-195	Sequence 195, App
7	372.5	30.9	181	2 US-08-478-572-195	Sequence 195, App
8	372.5	30.9	181	3 US-08-484-296-195	Sequence 195, App
9	354.5	29.4	181	2 US-08-482-142-197	Sequence 197, App
10	354.5	29.4	181	2 US-08-478-572-197	Sequence 197, App
11	354.5	29.4	181	3 US-08-484-296-197	Sequence 197, App
12	295	24.5	215	3 US-09-005-298-38	Sequence 38, Appl
13	295	24.5	215	4 US-08-768-619-38	Sequence 38, Appl
14	278	23.1	213	3 US-09-005-298-22	Sequence 22, Appl
15	278	23.1	213	4 US-08-768-619-22	Sequence 22, Appl
16	278	23.1	213	5 PCT-US96-09848-21	Sequence 21, Appl
17	278	23.1	216	1 US-08-546-712-2	Sequence 2, Appl
18	278	23.1	216	2 US-08-751-105-2	Sequence 3, Appl
19	254	21.1	220	4 US-09-462-846-3	Sequence 3, Appl
20	253	21.0	206	3 US-08-821-994-77	Sequence 77, Appl
21	252	20.9	220	3 US-09-004-731-95	Sequence 95, Appl
22	252	20.9	220	3 US-08-749-699-95	Sequence 95, Appl
23	252	20.9	220	4 US-09-004-729-95	Sequence 95, Appl
24	251.5	20.9	200	4 US-09-325-932A-150	Sequence 150, App
25	248	20.6	215	1 US-08-451-409A-1	Sequence 1, Appl
26	247	20.5	212	3 US-08-860-255A-4	Sequence 4, Appl
27	239.5	19.9	213	2 US-08-578-701A-1	Sequence 1, Appl

28 239.5 19.9 213 2 US-08-360-693-1 Sequence 1, Appl  
29 238 19.7 181 2 US-08-482-142-193 Sequence 193, App  
30 238 19.7 181 2 US-08-478-572-193 Sequence 193, App  
31 238 19.7 181 3 US-08-484-296-193 Sequence 193, App  
32 215.5 17.9 202 3 US-08-821-994-78 Sequence 78, Appl  
33 211 17.5 198 4 US-09-325-932A-151 Sequence 151, App  
34 208 17.2 129 4 US-09-325-932A-159 Sequence 159, App  
35 201 16.7 147 4 US-09-325-932A-173 Sequence 173, App  
36 197.5 16.4 160 4 US-09-325-932A-155 Sequence 155, App  
37 195 16.2 204 4 US-09-325-932A-145 Sequence 145, App  
38 181 15.0 130 4 US-09-325-932A-164 Sequence 164, App  
39 175.5 14.6 157 4 US-08-424-361B-17 Sequence 17, Appl  
40 174 14.4 168 3 US-08-821-994-45 Sequence 45, Appl  
41 173 14.3 29 2 US-08-482-142-39 Sequence 39, Appl  
42 173 14.3 29 2 US-08-478-572-39 Sequence 39, Appl  
43 173 14.3 29 3 US-08-484-296-39 Sequence 39, Appl  
44 172 14.3 119 4 US-09-325-932A-168 Sequence 168, App  
45 167 13.8 167 3 US-08-821-994-44 Sequence 44, Appl

## ALIGNMENTS

### RESULT 1

US-07-945-288-11

; Sequence 11, Application US/07945288

; Patent No. 5433948

; GENERAL INFORMATION:

; APPLICANT: Thomas, Wayne R.

; APPLICANT: Chua, Kaw-Yan

; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 STATE STREET, SUITE 510

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/945,288

; FILING DATE: 19920910

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 580,655

; FILING DATE: 11 SEPTEMBER 1990

; APPLICATION NUMBER: 458,642

; FILING DATE: 13 FEBRUARY 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: MANDRAGOURAS, AMY E.

; REGISTRATION NUMBER: P36,207

; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 222 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 50

; OTHER INFORMATION: /label-xaa is His or Tyr

; FEATURE:

; NAME/KEY: misc feature

```
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
; US-07-945-288-11

Query Match          97.6%; Score 1177; DB 1; Length 222;
Best Local Similarity 97.7%; Pred. No. 2.5e-126;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120
QY 121 PPNANKIREALQTHSAIAVIGIKDLPFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180
DB 121 PPNANKIREALQTHSAIAVIGIKDLPFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180
QY 181 QGVYWIIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222
DB 181 QGVYWIIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222

RESULT 2
US-08-462-831-11
; Sequence 11, Application US/08462831
; Patent No. 5552142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 459,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 50
; OTHER INFORMATION: /label=Xaa is His or Tyr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
; US-08-462-831-11

Query Match          97.6%; Score 1177; DB 1; Length 222;
Best Local Similarity 97.7%; Pred. No. 2.5e-126;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
DB 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120
QY 121 PPNANKIREALQTHSAIAVIGIKDLPFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180
DB 121 PPNANKIREALQTHSAIAVIGIKDLPFRHYDGRTHIIQRDNGYQPNYHVNIVGYSNA 180
QY 181 QGVYWIIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222
DB 181 QGVYWIIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVIL 222

RESULT 3
US-08-461-809-11
; Sequence 11, Application US/08461809
; Patent No. 5770202
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,809
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 50
; OTHER INFORMATION: /label=Xaa is His or Tyr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
;
; US-08-461-809-11
;
; Query Match 97.6%; Score 1177; DB 1; Length 222;
; Best Local Similarity 97.7%; Pred. No. 2.5e-126;
; Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
; DB 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
;
; QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNQAFGISNYCQIY 120
; DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNQAFGISNYCQIY 120
;
; QY 121 PPANKIREALQTHSAIAVLIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
; DB 121 PPANKIREALQTHSAIAVLIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
;
; QY 181 QGVDDYWIVRNSWDTNWDNGNGYGFPAANIDLMWIEEYPPVVIL 222
; DB 181 QGVDDYWIVRNSWDTNWDNGNGYGFPAANIDLMWIEEYPPVVIL 222
;
; RESULT 4
; US-08-461-441-11
; Sequence 11, Application US/08461441
; Patent No. 5773002
; GENERAL INFORMATION:
; APPLICANT:
;
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,441
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 50
; OTHER INFORMATION: /label=Xaa is His or Tyr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
;
; US-08-461-809-11
;
; Query Match 97.6%; Score 1177; DB 1; Length 222;
; Best Local Similarity 97.7%; Pred. No. 2.5e-126;
; Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
; DB 1 TNACSSINGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
;
; QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNQAFGISNYCQIY 120
; DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNQGVQESYRYVAREQSCRRPNQAFGISNYCQIY 120
;
; QY 121 PPANKIREALQTHSAIAVLIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
; DB 121 PPANKIREALQTHSAIAVLIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
;
; QY 181 QGVDDYWIVRNSWDTNWDNGNGYGFPAANIDLMWIEEYPPVVIL 222
; DB 181 QGVDDYWIVRNSWDTNWDNGNGYGFPAANIDLMWIEEYPPVVIL 222
;
; RESULT 4
; US-08-461-441-11
; Sequence 11, Application US/08461441
; Patent No. 5773002
; GENERAL INFORMATION:
; APPLICANT:
;
```

QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHNAVNIYGYSNA 180  
Db 121 PPNXNKIREALQTHXAIKDLDAFRHYDGRITIIQRDNGYQPNYHNAVNIYGYSNA 180  
QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMMEIEYPPYVIL 222  
Db 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMMEIEYPPYVIL 222

## RESULT 5

PCT-US93-08518-11  
; Sequence 11, Application PC/TUS9308518  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08518  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/945,288  
; FILING DATE: 10 SEPTEMBER 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRAGOURAS, AMY E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 222 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 50  
; OTHER INFORMATION: /label=Xaa is His or Tyr  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 81  
; OTHER INFORMATION: /label=Xaa is Glu or Lys  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 124  
; OTHER INFORMATION: /label=Xaa is Ala or Val  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 136  
; OTHER INFORMATION: /label=Xaa is Ser or Thr  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 215  
; OTHER INFORMATION: /label=Xaa is Glu or Gln  
PCT-US93-08518-11

Query Match 97.6%; Score 1177; DB 5; Length 222;  
Best Local Similarity 97.7%; Pred. No. 2.5e-126;

Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TNACSGINGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEO 60  
Db 1 TNACSGINGNAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAEO 60  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNQVQESYYRYVAREQSCRPNQORFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIXYIQHNGVQESYYRYVAREQSCRPNQORFGISNYCQIY 120  
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHNAVNIYGYSNA 180  
Db 121 PPNXNKIREALQTHXAIKDLDAFRHYDGRITIIQRDNGYQPNYHNAVNIYGYSNA 180  
QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMMEIEYPPYVIL 222  
Db 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMMEIEYPPYVIL 222

## RESULT 6

US-08-482-142-195  
; Sequence 195, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-Chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 195:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 181 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-142-195

Query Match 30.9%; Score 372.5; DB 2; Length 181;  
Best Local Similarity 62.6%; Pred. No. 2.2e-34;  
Matches 82; Conservative 8; Mismatches 14; Indels 27; Gaps 4;

QY 40 VAATESAYLAHRNQSLLDARQELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAR 99  
Db 24 VAATESAYLAYRNTSLDSEQLVD-----EYIQHNGVQVESYRYVAR 67  
QY 100 EQSCRRPNAORFGISNYCQIYPPNAN---KIREALAQTHSAJAVIIGIKDLDAFRHYDG 155  
Db 68 EQSCRRPNAQLE-----AVFEANQNTKTAKIE--IKASIDGLEVIIGIKDLDAFRHYDG 120  
QY 156 RTIIQRDNGYQ 166  
Db 121 RTIIQRDNGYQ 131

RESULT 7  
US-08-478-572-195  
; Sequence 195, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 195:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 181 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-478-572-195

Query Match 30.9%; Score 372.5; DB 2; Length 181;  
Best Local Similarity 62.6%; Pred. No. 2.2e-34;  
Matches 82; Conservative 8; Mismatches 14; Indels 27; Gaps 4;

QY 40 VAATESAYLAHRNQSLLDARQELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAR 99  
Db 24 VAATESAYLAYRNTSLDSEQLVD-----EYIQHNGVQVESYRYVAR 67  
QY 100 EQSCRRPNAORFGISNYCQIYPPNAN---KIREALAQTHSAJAVIIGIKDLDAFRHYDG 155  
Db 68 EQSCRRPNAQLE-----AVFEANQNTKTAKIE--IKASIDGLEVIIGIKDLDAFRHYDG 120

Db 68 EQSCRRPNAQLE-----AVFEANQNTKTAKIE--IKASIDGLEVIIGIKDLDAFRHYDG 120  
QY 156 RTIIQRDNGYQ 166  
Db 121 RTIIQRDNGYQ 131

RESULT 8  
US-08-484-296-195  
; Sequence 195, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,296  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 195:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 181 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-484-296-195

Query Match 30.9%; Score 372.5; DB 3; Length 181;  
Best Local Similarity 62.6%; Pred. No. 2.2e-34;  
Matches 82; Conservative 8; Mismatches 14; Indels 27; Gaps 4;

QY 40 VAATESAYLAHRNQSLLDARQELVDCASQHGCHGDTIPRGIEYIQHNGVQVESYRYVAR 99  
Db 24 VAATESAYLAYRNTSLDSEQLVD-----EYIQHNGVQVESYRYVAR 67  
QY 100 EQSCRRPNAORFGISNYCQIYPPNAN---KIREALAQTHSAJAVIIGIKDLDAFRHYDG 155  
Db 68 EQSCRRPNAQLE-----AVFEANQNTKTAKIE--IKASIDGLEVIIGIKDLDAFRHYDG 120  
QY 156 RTIIQRDNGYQ 166  
Db 121 RTIIQRDNGYQ 131





```
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/768,619
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/486,036
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Connell, Gary J.
/ REGISTRATION NUMBER: 32,020
/ REFERENCE/DOCKET NUMBER: 2618-33-C1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (303) 863-9700
/ TELEFAX: (303) 863-0223
/ INFORMATION FOR SEQ ID NO: 38:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 215 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-768-619-38

Query Match      24.5%; Score 295; DB 4; Length 215;
Best Local Similarity 35.8%; Pred. No. 2.2e-25;
Matches 73; Conservative 30; Mismatches 89; Indels 12; Gaps 7;

QY 11 PAEIDLQMRVTPIRMQGGSCSWAFSGVAATESAYLAHRNOSLDLAEQELVDCA---S 67
DB 2 PKYVDRKRGGYVTPAKEQGLCGSCYAFASAAAALEAYNKKTKNKLKLDLSPQNLDCDCTWDLG 61
QY 68 QHCGHGDTPRGIEYIQHNGVVOESYRYV--AREQSCRRPNAQRFGISNYCOIYPPNAN 125
DB 62 NNGCHGGFNPAPFYASKAGIASAKYPVHTARTCTYWRKDIVAATDNGYTRIQQGDEK 121
QY 126 KIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNITVGY--SNAQGV 183
DB 122 GLQYAVAKGFPVVVGISGQ--HDFKFKYS-GVYSSDQCRVFN-HAVLVVGYGTSKKHG- 176
QY 184 DYWIVRNSWDTNWGDNGYGYFAAN 207
DB 177 DYWIINKSWGNTWNGRNGYGYMKRN 200

RESULT 14
US-09-005-298-22
/ Sequence 22, Application US/09005298
/ Patent No. 6365392
/ GENERAL INFORMATION:
/ APPLICANT: Tripp, Cynthia A.
/ APPLICANT: Wisniewski, Nancy
/ APPLICANT: Grieve, Robert B.
/ APPLICANT: Frank, Glenn R.
/ TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sheridan Ross P.C.
/ STREET: 1700 Lincoln Street, Suite 3500
/ CITY: Denver
/ STATE: Colorado
/ COUNTRY: U.S.A.
/ ZIP: 80203
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/005,298
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/486,036
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Connell, Gary J.
/ REGISTRATION NUMBER: 32,020
```

```
/ APPLICATION NUMBER: 08/768,619
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Connell, Gary J.
/ REGISTRATION NUMBER: 32,020
/ REFERENCE/DOCKET NUMBER: 2618-33-C1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (303) 863-9700
/ TELEFAX: (303) 863-0223
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 213 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-005-298-22

Query Match      23.1%; Score 278; DB 3; Length 213;
Best Local Similarity 34.8%; Pred. No. 1.9e-23;
Matches 71; Conservative 27; Mismatches 92; Indels 14; Gaps 8;

QY 11 PAEIDLQMRVTPIRMQGGSCSWAFSGVAATESAYLAHRNOSLDLAEQELVDCA---S 67
DB 2 PKYVDRKRGGYVTPAKEQGLCGSCYAFASCSIRSLIYKTKNKLKLDLSPQNLDCDCTWDLG 61
QY 68 QHCGHGDTPRGIEYIQHNGVVOESYRYV--AREQSCRRPNAQRFGISNYCOIYPPNAN 125
DB 62 NNGCHGGFNPAPFYASKAGIASAKYPVHTARTCTYWRKDIVAATDNGYTRIQQGD-E 120
QY 126 KIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNITVGY--SNAQGV- 183
DB 121 KGLNLMQ----LTVVVGISGQHDFFKFKYS-GVYSSDQCRVFN-HAVLVVGYGTSQKTR 174
QY 184 DYWIVRNSWDTNWGDNGYGYFAAN 207
DB 175 DYWIINKSWGNTWNGRNGYGYMKRN 198

RESULT 15
US-08-768-619-22
/ Sequence 22, Application US/08768619
/ Patent No. 6419923
/ GENERAL INFORMATION:
/ APPLICANT: Tripp, Cynthia A.
/ APPLICANT: Wisniewski, Nancy
/ APPLICANT: Grieve, Robert B.
/ APPLICANT: Frank, Glenn R.
/ TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sheridan Ross P.C.
/ STREET: 1700 Lincoln Street, Suite 3500
/ CITY: Denver
/ STATE: Colorado
/ COUNTRY: U.S.A.
/ ZIP: 80203
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/768,619
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/486,036
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Connell, Gary J.
/ REGISTRATION NUMBER: 32,020
```



```
; REFERENCE/DOCKET NUMBER: 2618-33-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-768-619-22

Query Match      23.1%; Score 278; DB 4; Length 213;
Best Local Similarity 34.8%; Pred. No. 1.9e-23;
Matches 71; Conservative 27; Mismatches 92; Indels 14; Gaps 8;

QY 11 PAETDLQMRVTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDCA---S 67
Db 2 PKYDWRKRGYVTPAKEQGLCGSCYAFCSCSIRSLIYKTKNKLDDLPQNILDCTWDLG 61

QY 68 QHGCHGDTIPRGIEYIQHNGVVQESYYRYV--AREQSCRPNRQAFGISNYCQIYPNAN 125
Db 62 NNGCHGGFMNPAFYASKAGIASEAKYPYVHTARRTCYWRKDIVAATDNGYTRIQQGD-E 120

QY 126 KIREALAQTHSAIAVIIGIKOLD-APRHYDGRITIIQRDNGYQPNYHAWNIVGYSNAQGV- 183
Db 121 KGLNMLWQ----LTVVVGISGYQHDFFKFKYS-GVYSSDQCRVPN-HAVLVVVGYSOKTR 174

QY 184 DYWIVRNSWDITNMGDNGYGYFAAN 207
Db 175 DYWIIRKNSWGINWNRNGYGYMKRN 198
```

Search completed: May 19, 2005, 17:36:14  
Job time : 24 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:50:19 ; Search time 12 Seconds  
(without alignments)  
80.181 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	34.4	10	2 S70336	napin small chain
2	20	32.8	7	2 A12016	formylglycinamide
3	19	31.1	9	2 I50633	c-rel protein - ch
4	19	31.1	10	2 D61440	polysacacturonase
5	19	31.1	10	2 S27873	hypothetical prote
6	18	29.5	9	1 YFPG	thymic factor - pi
7	18	29.5	9	2 A60957	thymocyte growth p
8	18	29.5	10	2 A61289	streptopain (EC 3.
9	17	27.9	3	3 A23751	spinal cord peptid
10	17	27.9	6	2 PT0727	i-cell receptor be
11	17	27.9	10	2 S06964	hypothetical prote
12	17	27.9	10	2 A60527	sperm-activating p
13	17	27.9	10	2 H60787	sperm-activating p
14	17	27.9	10	2 F60787	sperm-activating p
15	17	27.9	10	2 A60787	sperm-activating p
16	17	27.9	10	2 D60787	sperm-activating p
17	17	27.9	10	2 B60787	sperm-activating p
18	17	27.9	10	2 D60588	sperm-activating p
19	17	27.9	10	2 B60588	sperm-activating p
20	17	27.9	10	2 C60588	sperm-activating p
21	17	27.9	10	2 F60527	sperm-activating p
22	17	27.9	10	2 A60588	sperm-activating p
23	17	27.9	10	2 A60788	sperm-activating p
24	17	27.9	10	2 D60527	sperm-activating p
25	17	27.9	10	2 C39572	sperm-activating p
26	17	27.9	10	2 F60572	sperm-activating p
27	17	27.9	10	2 C60527	sperm-activating p
28	17	27.9	10	2 E60527	sperm-activating p
29	17	27.9	10	2 G60527	sperm-activating p

30	17	27.9	10	2 E39572	sperm-activating p
31	17	27.9	10	2 D60788	sperm-activating p
32	17	27.9	10	2 E60788	sperm-activating p
33	17	27.9	10	2 C60788	sperm-activating p
34	17	27.9	10	2 F60589	sperm-activating p
35	17	27.9	10	2 C60589	sperm-activating p
36	17	27.9	10	2 D60589	sperm-activating p
37	17	27.9	10	2 F60588	sperm-activating p
38	17	27.9	10	2 B60589	sperm-activating p
39	16	26.2	5	2 F22565	R-phycoerythrin ga
40	16	26.2	6	4 I79564	hypothetical TCU3
41	16	26.2	10	2 S51912	hemagglutinin - in
42	16	26.2	10	2 A35356	hypothetical prote
43	16	26.2	10	2 B61440	polysacacturonase
44	15	24.6	4	2 S47552	ubiquitin - rat
45	15	24.6	5	2 A33882	cadmium-binding pe

ALIGNMENTS

RESULT 1

S70336  
napin small chain S1A - Swedish turnip (fragment)  
C;Species: Brassica napus var. rapifera (Swedish turnip, rutabaga)  
C;Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004  
C;Accession: S70336  
R;Neumann, G.M.; Condron, R.; Thomas, I.; Polya, G.M.  
Biochim. Biophys. Acta 1295, 23-33, 1996  
A;Title: Purification and sequencing of multiple forms of Brassica napus seed napin small chain  
A;Reference number: S70336; MUID:96283790; PMID:8679670  
A;Accession: S70336  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <NEU>  
A;Cross-references: UNIPROT:Q42469; UNIPROT:Q9S9F2

Query Match 34.4%; Score 21; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MQSGCG 7  
Db 3 MQSGSG 8

RESULT 2

Al2016  
formylglycinamide ribonucleotide amidotransferase (EC 2.---) - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 13-Mar-1997  
C;Accession: A12016; B12016  
R;Ohnoki, S.; Hong, B.S.; Buchanan, J.M.  
Fed. Proc. 35, 1549, 1976  
A;Title: Amino acid sequence at glutamine active site for FGAR-amidotransferase.  
A;Reference number: A91459  
A;Accession: A12016  
A;Molecule type: protein  
A;Residues: 1-7 <OHN>  
A;Experimental source: liver, peptide 1  
A;Accession: B12016  
A;Molecule type: protein  
A;Residues: 1-5 <OH2>  
A;Experimental source: liver, peptide 2  
C;Keywords: transferase

Query Match 32.8%; Score 20; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GGCGSC 9  
Db 1 GVCDBC 6

RESULT 3  
 I50633  
 c-rel protein - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C:Accession: I50633  
 R:Hannink, M.; Temin, H.M.  
 Oncogene 5, 1843-1850, 1990  
 A:Title: Structure and autoregulation of the c-rel promoter.  
 A:Reference number: I50633; MUID:91133738; PMID:2284104  
 A:Accession: I50633  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-9 <HAN>  
 A:Cross-references: UNIPROT:Q92009; EMBL:X56440; NID:g63338; PIDN:CAA39822.1; PID:g58448  
  
 Query Match 31.1%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 MQGGCG 7  
 : |||  
 Db 3 VSGGAG 8  
  
 RESULT 4  
 D61440  
 polygalacturonase (EC 3.2.1.15) IV - Aspergillus sp. (fragment)  
 C:Species: Aspergillus sp.  
 C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004  
 C:Accession: D61440  
 R:Stratilova, E.; Markovic, O.; Skrovinova, D.; Rexova-Benkova, L.; Jorvall, H.  
 J. Protein Chem. 12, 15-22, 1993  
 A:Title: Pectinase Aspergillus sp. polygalacturonase: multiplicity, divergence, and structure.  
 A:Reference number: A61440; MUID:93151962; PMID:8427629  
 A:Accession: D61440  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <STR>  
 A:Cross-references: UNIPROT:Q7M501  
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
  
 Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7e+03; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 7 GSC 9  
 : |||  
 Db 1 GSC 3  
  
 RESULT 5  
 S27873  
 hypothetical protein 2 LRH-1 5'-region - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Aug-2004  
 C:Accession: S27873  
 R:Tugwood, J.D.; Issemann, I.; Green, S.  
 submitted to the EMBL Data Library, February 1992  
 A:Description: LRH-1: A nuclear hormone receptor active in the absence of exogenous ligand.  
 A:Reference number: S27873  
 A:Accession: S27873  
 A:Molecule type: mRNA  
 A:Residues: 1-10 <TUG>  
 A:Cross-references: UNIPROT:Q61807; EMBL:M81385; NID:g198872; PIDN:AAA39446.1; PID:g198872  
  
 Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7e+03; Indels 0; Gaps 0;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 8 SCN 10

Db 3 SCN 5  
  
 RESULT 6  
 YFPG  
 thymic factor - pig  
 N:Alternate names: FTS (facteur thymique serieque)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 09-Jul-2004  
 C:Accession: A01523; A60983  
 R:Pleau, J.M.; Dardenne, M.; Blouquit, Y.; Bach, J.F.  
 J. Biol. Chem. 252, 8045-8047, 1977  
 A:Title: Structural study of circulating thymic factor: a peptide isolated from pig serum.  
 A:Reference number: A01523; MUID:78026571; PMID:914862  
 A:Accession: A01523  
 A:Molecule type: protein  
 A:Residues: 1-9 <PLE>  
 A:Cross-references: UNIPROT:P01255  
 R:Bach, J.F.; Dardenne, M.; Pleau, J.M.; Rosa, J.  
 Nature 266, 55-57, 1977  
 A:Title: Biochemical characterisation of a serum thymic factor.  
 A:Reference number: A60983; MUID:77123829; PMID:300146  
 A:Accession: A60983  
 A:Molecule type: protein  
 A:Residues: 'Z', '2-4', 'Z', '6-9 <BAC>  
 A:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunoassays.  
 C:Comment: See PIR:A60957 (sheep) for discussion of another possible N-terminal modification.  
 C:Superfamily: thymic factor  
 C:Keywords: pyrrolidone carboxylic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
  
 Query Match 29.5%; Score 18; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 RMQGG 5  
 : |||  
 Db 3 KSQGG 7  
  
 RESULT 7  
 A60957  
 thymocyte growth peptide - sheep  
 N:Contains: FTS (facteur thymique serieque)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A60957  
 R:Ernstroem, U.; Gafvelin, G.; Rudja, J.M.  
 Biosci. Rep. 10, 403-412, 1990  
 A:Title: Purification of thymocyte growth peptide (TGP) from sheep thymus. Relationship to  
 A:Reference number: A60957; MUID:91064427; PMID:2249004  
 A:Accession: A60957  
 A:Molecule type: protein  
 A:Residues: 1-9 <ERN>  
 A:Cross-references: UNIPROT:Q7M3C5  
 A:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunoassays.  
 C:Comment: This peptide was isolated in two forms. One form contained the pyrrolidone carboxylic acid (Glx) (in thymocyte growth peptide) contains a large, non-peptide blocking group with a high molecular weight.  
 C:Superfamily: thymic factor  
 C:Keywords: blocked amino end; pyrrolidone carboxylic acid (Glx) (in FTS) #status experimental  
 F:1/Modified site: pyrrolidone carboxylic acid (Glx) (in thymocyte growth peptide) #status experimental  
 F:1/Modified site: blocked amino end (Glx) (in thymocyte growth peptide) #status experimental  
  
 Query Match 29.5%; Score 18; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 RMQGG 5  
 : |||  
 Db 3 KSQGG 7



sperm-activating peptide (Thr-5 speract) - sea urchin (Anthocidaris crassispina)  
C:Species: Anthocidaris crassispina  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 16-Aug-2004  
C:Accession: A60787  
R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
Comp. Biochem. Physiol. B 89, 687-693, 1988  
A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
A:Reference number: A60787; PMID:88242184; PMID:3378407  
A:Accession: A60787  
A:Molecule type: protein  
A:Residues: 1-10 <SUZ>  
A:Cross-references: UNIPROT:Q7M4D6  
C:Comment: This oligopeptide from egg jelly is one of several from this species, all of  
at shows some, but not absolute, species restriction.

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. NO. 1.4e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCG 7  
DB 7 GGVG 10

## RESULT 14

F60787  
sperm-activating peptide (Ser-5 speract) - sea urchin (Anthocidaris crassispina)  
C:Species: Anthocidaris crassispina  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 16-Aug-2004  
C:Accession: F60787  
R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
Comp. Biochem. Physiol. B 89, 687-693, 1988  
A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
A:Reference number: A60787; PMID:88242184; PMID:3378407  
A:Accession: F60787  
A:Molecule type: protein  
A:Residues: 1-10 <SUZ>  
A:Cross-references: UNIPROT:Q7M4D7  
C:Comment: This oligopeptide from egg jelly is one of several from this species, all of  
at shows some, but not absolute, species restriction.

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. NO. 1.4e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCG 7  
DB 7 GGVG 10

## RESULT 15

A60787  
sperm-activating peptide (speract) - sea urchin (Hemicentrotus pulcherrimus)  
C:Species: Hemicentrotus pulcherrimus  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 09-Jul-2004  
C:Accession: A60787  
R:Suzuki, N.; Kajiwara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
Comp. Biochem. Physiol. B 89, 687-693, 1988  
A:Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
A:Reference number: A60787; PMID:88242184; PMID:3378407  
A:Accession: A60787  
A:Molecule type: protein  
A:Residues: 1-10 <SUZ>  
A:Cross-references: UNIPROT:Q25121  
C:Comment: This oligopeptide from egg jelly is one of several from this species, all of  
at shows some, but not absolute, species restriction.  
C:Superfamily: unassigned animal peptides

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. NO. 1.4e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCG 7

DB 7 GGVG 10

Search completed: May 19, 2005, 17:59:33  
Job time : 16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:54:20 ; Search time 111 Seconds  
(without alignments)  
46.133 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMGGGSGSN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 2548

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	47.5	10	Q6X2S9	Q6x2s9 homo sapien
2	28	45.9	10	Q8WP04	Q8wp04 ateles belz
3	26	42.6	9	O12096	O12096 caprine art
4	26	42.6	9	O12098	O12098 caprine art
5	26	42.6	9	O12100	O12100 caprine art
6	26	42.6	9	O12102	O12102 caprine art
7	26	42.6	9	O12104	O12104 caprine art
8	21	34.4	10	P82938	P82938 hordeum vul
9	20	32.8	10	Q9X3M2	Q9x3m2 prochloroco
10	19	31.1	9	Q92009	Q92009 gallus gall
11	19	31.1	10	O7M501	O7m501 aspergillus
12	19	31.1	10	O8NEY9	O8ney9 homo sapien
13	19	31.1	10	Q96QA7	Q96qa7 homo sapien
14	19	31.1	10	O61807	O61807 mus musculu
15	19	31.1	10	O8UVM2	O8uvw2 oreochromis
16	18	29.5	8	O15890	O15890 homo sapien
17	18	29.5	9	THYF_PIG	P01255 sus scrofa
18	18	29.5	9	Q9H3F3	Q9h3f3 homo sapien
19	18	29.5	9	Q7M3C5	Q7m3c5 ovis aries
20	18	29.5	9	O99JF4	O99jf4 mus musculu
21	18	29.5	10	O7M0W1	O7m0w1 streptococc
22	18	29.5	10	O64G36	O64g36 avian infec
23	17	27.9	7	ASCL_ALLAS	P84071 allium asca
24	17	27.9	8	Q9Y4X6	Q9y4x6 homo sapien
25	17	27.9	10	O71VB8	O71vb8 homo sapien
26	17	27.9	10	Q7M3T4	Q7m3t4 tripneustes
27	17	27.9	10	O7M3T5	O7m3t5 tripneustes
28	17	27.9	10	O7M3T6	O7m3t6 tripneustes
29	17	27.9	10	O7M3T7	O7m3t7 tripneustes
30	17	27.9	10	O7M3T8	O7m3t8 tripneustes
31	17	27.9	10	O7M3T9	O7m3t9 tripneustes

32	17	27.9	10	2	Q7M3U0	Q7m3u0 tripneustes
33	17	27.9	10	2	Q7M4B2	Q7m4b2 stronglyloce
34	17	27.9	10	2	Q7M4B3	Q7m4b3 stronglyloce
35	17	27.9	10	2	Q7M4B4	Q7m4b4 stronglyloce
36	17	27.9	10	2	Q7M4B5	Q7m4b5 stronglyloce
37	17	27.9	10	2	Q7M4B6	Q7m4b6 stronglyloce
38	17	27.9	10	2	Q7M4B9	Q7m4b9 heterocentr
39	17	27.9	10	2	Q7M4C0	Q7m4c0 heterocentr
40	17	27.9	10	2	Q7M4C3	Q7m4c3 echinometra
41	17	27.9	10	2	Q7M4C4	Q7m4c4 echinometra
42	17	27.9	10	2	Q7M4C5	Q7m4c5 echinometra
43	17	27.9	10	2	Q7M4C8	Q7m4c8 stronglyloce
44	17	27.9	10	2	Q7M4D1	Q7m4d1 hemiceutrot
45	17	27.9	10	2	Q7M4D2	Q7m4d2 pseudocentr

ALIGNMENTS

```
RESULT 1
Q6X2S9
ID Q6X2S9 PRELIMINARY; PRT; 10 AA.
AC Q6X2S9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Translocase of the inner mitochondrial membrane (Fragment).
GN Names=TIM23;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin.
RX MEDLINE=22890074; PubMed=14527731; DOI=10.1016/S0378-1119(03)00738-8;
RA Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.;
RT "Human poly(ADP-ribose) glycohydrolase (PARG) gene and the common
RT promoter sequence it shares with inner mitochondrial membrane
RT translocase 23 (TIM23).";
RL Gene 314:181-190(2003).
DR ENBL; AY258588; AAP83315.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;

Query Match 47.5%; Score 29; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MQGGGSGSN 10
|:|:|
Db 1 MEGGGGSGN 9

RESULT 2
Q8WP04
ID Q8WP04 PRELIMINARY; PRT; 10 AA.
AC Q8WP04;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Oculocutaneous albinism type II (Fragment).
GN Names=OCA2;
OS Ateles belzebuth chamek (Chamek spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=118643;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21636947; PubMed=11778696; DOI=10.1023/A:1012900206671;
RA Seanez H.N., Lima C.R., Lemos B., Bonvicino C.R., Moreira M.A.M.,
RA Canavez F.C.;
RT "Gene assignment in Ateles paniscus chamek (Platyrrhini, Primates).
```

RT Allocation of 18 markers of human syntenic groups 1,2,7,14,15,17 and 22,";  
 RL Chromosome Res. 9:631-639(2001).  
 DR EMBL; AF375652; AAL31489.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 901 MW; 22DF477DD87EA5B8 CRC64;

Query Match 45.9%; Score 28; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+02; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0;

QY 6 CGSC 9  
 ||||  
 Db 2 CGSC 5

RESULT 3  
 OI2096  
 ID OI2096 PRELIMINARY; PRT; 9 AA.  
 AC OI2096;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81439; AAB60832.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 4  
 OI2098  
 ID OI2098 PRELIMINARY; PRT; 9 AA.  
 AC OI2098;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81440; AAB60835.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 5  
 OI2100  
 ID OI2100 PRELIMINARY; PRT; 9 AA.  
 AC OI2100;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81441; AAB60836.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 6  
 OI2102  
 ID OI2102 PRELIMINARY; PRT; 9 AA.  
 AC OI2102;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 7  
 OI2104  
 ID OI2104 PRELIMINARY; PRT; 9 AA.  
 AC OI2104;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81441; AAB60836.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 8  
 OI2106  
 ID OI2106 PRELIMINARY; PRT; 9 AA.  
 AC OI2106;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81441; AAB60836.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 9  
 OI2110  
 ID OI2110 PRELIMINARY; PRT; 9 AA.  
 AC OI2110;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 10  
 OI2114  
 ID OI2114 PRELIMINARY; PRT; 9 AA.  
 AC OI2114;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 11  
 OI2118  
 ID OI2118 PRELIMINARY; PRT; 9 AA.  
 AC OI2118;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 12  
 OI2122  
 ID OI2122 PRELIMINARY; PRT; 9 AA.  
 AC OI2122;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 ||||  
 Db 2 GCRLCN 7

RESULT 13  
 OI2126  
 ID OI2126 PRELIMINARY; PRT; 9 AA.  
 AC OI2126;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U81442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



```

RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81443; AAB60840.1; -.
FT NON_TER
SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.6e+06;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10
DB 2 GCRLCN 7

RESULT 8
P82938 ID P82938 PRELIMINARY; PRT; 10 AA.
AC P82938
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Unknown endosperm protein C (Fragment).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. BOMI; TISSUE=Starchy endosperm;
RX MEDLINE=21088911; PubMed=11271488;
RX DOI=10.1002/1522-2683(200011)21:17<3693::AID-ELPS3693>3.0.CO;2-I;
RA Kristoffersen H.E., Flengrød R.;
RT "Separation and characterization of basic barley seed proteins.";
RL Electrophoresis 21:3693-3700(2000)
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 8.5-9.0, ITS MW IS: 11.9 KDA.
FT NON_TER
FT NON_TER 10
SQ SEQUENCE 10 AA; 1053 MW; 9E562DC40AA87AAE CRC64;

Query Match 34.4%; Score 21; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 9.8e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGC 6
DB 1 GGC 3

RESULT 9
Q9X3M2 ID Q9X3M2 PRELIMINARY; PRT; 10 AA.
AC Q9X3M2
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Cytochrome b (Fragment).
GN Names:petB;
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070219; AAD23269.1; -.
FT NON_TER
FT NON_TER 10
SQ SEQUENCE 10 AA; 1076 MW; 75CA5CB05866D324 CRC64;

Query Match 32.8%; Score 20; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. No. 1.4e+04;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RMOGCGCS 8
DB 3 RKQGISGS 10

RESULT 10
Q92009 ID Q92009 PRELIMINARY; PRT; 9 AA.
AC Q92009
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE C-rel protein (p68-c-rel) (Fragment).
GN Names=c-rel proto-oncogene;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91133738; PubMed=2284104;
RA Hannink M., Temin H.M.;
RT "Structure and autoregulation of the c-rel promoter.";
RL Oncogene 5:1843-1850(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Hannink M., Temin H.M.;
RT "Structure and auto regulation of the c-rel promoter.";
RL Oncogene 0:0-0(1990).
DR EMBL; X56440; CAA39822.1; -.
DR EMBL; X56515; CAA39866.1; -.
DR PIR; I50633; I50633. 9
FT NON_TER
FT NON_TER 9
SQ SEQUENCE 9 AA; 805 MW; DE317DD87865A2CD CRC64;

Query Match 31.1%; Score 19; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MQGCGC 7
DB 3 VSGGAG 8

RESULT 11
Q7M501 ID Q7M501 PRELIMINARY; PRT; 10 AA.
AC Q7M501
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polygalacturonase (EC 3.2.1.15) IV (Fragment).
OS Aspergillus sp.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5065;
RN [1]
RP SEQUENCE.
RX MEDLINE=93151962; PubMed=8427629;
RA Stratilova E., Markovic O., Skrovinova D., Rexova-Benkova L.,
RA Jernvall H.;
RT "Pectinase Aspergillus sp. polygalacturonase: multiplicity,
RT divergence, and structural patterns linking fungal, bacterial, and
RT plant polygalacturonases.";
RL J. Protein Chem. 12:15-22(1993).
DR PIR; D61440; D61440.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
FT NON_TER
FT NON_TER 1

```

```

FT NON TER      10      10
SQ SEQUENCE    10 AA; 959 MW; 845236C5A1A9D1AE CRC64;

Query Match      31.1%; Score 19; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 GSC 9
      |||
Db      1 GSC 3

RESULT 12
ID Q8NEY9 PRELIMINARY; PRT; 10 AA.
AC Q8NEY9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Type II hair-specific keratin (Type II hair keratin) (Fragment).
GN Name=KRTHB6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bairwa N.K., Bamezai R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bairwa N.K., Bamezai R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037552; AK68688.1; -.
DR EMBL; AY203963; AA063472.1; -.
DR GO; GO:0005882; C:intermediate filament; IEA.
KW Keratin.
FT NON TER      10      10
SQ SEQUENCE    10 AA; 1034 MW; 9B53417EAB45B87E CRC64;

Query Match      31.1%; Score 19; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 CGS 8
      |||
Db      3 CGS 5

RESULT 13
ID Q96QA7 PRELIMINARY; PRT; 10 AA.
AC Q96QA7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BA15A2.1 (Cdc42 guanine exchange factor (GEF) 9 (Collybistin, PEM-2,
DE HPBM-2, KIAA0424) (Fragment).
GN Name=ARHGEF9;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL451106; CAC89408.1; -.
FT NON TER      10      10
SQ SEQUENCE    10 AA; 1122 MW; 39925CEB78640043 CRC64;

Query Match      31.1%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 2e+04;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 SCN 10
      |||
Db      3 SCN 5

us-09-867-159a-3.closed.rup

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 MOGGCG 7
      :|||
Db      4 IRGGSG 9

RESULT 14
ID Q61807 PRELIMINARY; PRT; 10 AA.
AC Q61807;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE LRH-1 protein.
GN Name=Nr5a2; Synonyms=LRH-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Alpk:APICD-1; TISSUE=Liver;
RX MEDLINE=22755858; PubMed=12672674;
RA Liu D.L., Liu W.Z., Li Q.L., Wang H.M., Qian D., Treuter E., Zhu C.;
RT "Expression and functional analysis of liver receptor homologue-1 as a
RL potential steroidogenic factor in rat ovary.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Alpk:APICD-1; TISSUE=Liver;
RA Tugwood J.D., Issemann I., Green S.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M81385; AAA39446.1; -.
DR PIR; S27873; S27873.
DR MGD; MGI:1346834; Nr5a2.
DR GO; GO:0003677; F:DNA binding; IDA.
DR GO; GO:0008206; P:bile acid metabolism; IMP.
DR GO; GO:0042632; P:cholesterol homeostasis; IMP.
SQ SEQUENCE    10 AA; 1133 MW; 998B68F5B7244EA5 CRC64;

Query Match      31.1%; Score 19; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 SCN 10
      |||
Db      3 SCN 5

RESULT 15
ID Q8UWV2 PRELIMINARY; PRT; 10 AA.
AC Q8UWV2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE AMPA receptor subunit 2 alpha (Fragment).
GN Name=Glur2a;
OS Oreochromis mossambicus (Mozambique tilapia) (Tilapia mossambica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8127;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21617183; PubMed=11741603; DOI=10.1016/S0014-5793(01)03183-0;
RA Kung S.-S., Chen Y.-C., Lin W.-H., Chen C.-C., Chow W.-Y.;
RT "Q/R RNA editing of the AMPA receptor subunit 2 (GRIA2) transcript
RT involves no later than the appearance of cartilaginous fishes.";
RL FEBS Lett. 509:277-281(2001).
DR EMBL; AF350050; AAL57192.1; -.

```

DR GO:0004872; F:receptor activity; IEA.

KW Receptor.

FT NON\_TER 1

FT NON\_TER 10

SQ SEQUENCE 10 AA; 1153 MW; 721A6E5AEB866C4 CRC64;

Query Match: 31.1%; Score 19; DB 2; Length 10;

Best Local Similarity 60.0%; Pred. No. 2e+04;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MQGQC 6

Db 2 MRQGC 6

Search completed: May 19, 2005, 18:15:42

Job time : 114 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:47:19 ; Search time 141 Seconds  
(without alignments)  
27.430 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 465227

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	10	5 AAO20569	Cysteine
2	61	100.0	10	5 ABB51909	Cysteine
3	37	60.7	8	2 AAR77350	Cysteine
4	36	59.0	9	2 ADI11045	Somatosta
5	36	59.0	9	2 ADI11044	Somatosta
6	36	59.0	9	6 ABUS5751	Polylinke
7	36	59.0	9	2 ADF69793	M. tuberc
8	36	59.0	10	2 ADI11047	Somatosta
9	32	52.5	10	4 AAG86614	Saccharom
10	31	50.8	8	2 AAW41566	Human cal
11	31	50.8	8	4 AAB97506	Peptide n
12	30	49.2	6	5 ABB08544	Human HCC
13	30	49.2	6	6 ABUS5750	Polylinke
14	30	49.2	6	7 ADF69792	M. tuberc
15	30	49.2	8	4 AAB97503	Peptide n
16	30	49.2	10	2 AAR69298	Gp IIB/II
17	30	49.2	10	2 ADD67499	Specific-
18	30	49.2	10	2 AAW60340	Tumour ho
19	30	49.2	10	2 AAW50579	GPIIb/III
20	30	49.2	10	2 AAW93666	Human bre
21	30	49.2	10	2 ADE25476	TC-99m la
22	30	49.2	10	3 AAY54962	Peptide 1
23	30	49.2	10	3 AAY95503	GPIIb/III
24	30	49.2	10	3 AAB21757	Human bre
25	30	49.2	10	4 AAE06335	Human bre

26	30	49.2	10	5 ABB51909	Abb51909 Human 34P
27	30	49.2	10	5 ABB51815	Abb51815 Human 34P
28	30	49.2	10	5 ABB51593	Abb51593 Human 34P
29	30	49.2	10	5 ABB51491	Abb51491 Human 34P
30	30	49.2	10	5 ABB51705	Abb51705 Human 34P
31	29	47.5	8	2 AAW52097	AAW52097 Targettin
32	29	47.5	8	2 AAY27435	AAY27435 Rat HICP
33	29	47.5	8	5 AAO20971	AAO20971 8-mer ins
34	29	47.5	8	5 AAB15661	AAB15661 Insulin-1
35	29	47.5	8	6 ABP56100	ABP56100 Human IGF
36	29	47.5	8	7 ADH02887	ADH02887 AscII pep
37	29	47.5	8	7 ADH02888	ADH02888 Sali pep
38	29	47.5	9	2 AAW06390	AAW06390 Cyclic pe
39	29	47.5	9	2 ADE25470	ADE25470 GPIIb/III
40	29	47.5	9	2 ADE25444	ADE25444 GPIIb/III
41	29	47.5	9	2 ADH59029	ADH59029 GFIIb/IIa
42	29	47.5	9	3 AAY54928	AAY54928 Peptide 1
43	29	47.5	9	4 AAG88625	AAG88625 HER2/NEU
44	29	47.5	9	8 ADE64416	ADE64416 Radiophar
45	29	47.5	9	8 ADH58651	ADH58651 Radiophar

#### ALIGNMENTS

##### RESULT 1

AAO20569

ID AAO20569 standard; peptide; 10 AA.

XX AC AAO20569;

XX DT 02-JAN-2003 (first entry)

XX DE Cysteine protease epitope peptide region, SEQ ID No 3.

XX KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
XX KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
XX KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
XX KW atopic eczema; epitope.

XX OS Dermatophagoides pteronyssinus.

XX PN WO200278736-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-FR001098.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PR 03-MAY-2001; 2001FR-00005929.

XX PR 29-MAY-2001; 2001US-00867159.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-750636/81.

XX DR Antiallergic compositions containing an anti-histamine, a histamine  
XX PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
XX PT for the allergen.

XX PS Claim 14; Page 11; 32pp; French.

XX CC The invention relates to antiallergic compositions containing an anti-  
XX CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
XX CC isolated nucleic acid molecule that has at least one polynucleotide  
XX CC sequence coding for the allergen, together with a pharmaceutical carrier.  
XX CC The pharmaceutical composition of the invention is useful as a non-  
XX CC specific antiallergic treatment, and also useful in the treatment of  
XX CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
XX CC allergic and atopic eczema. This sequence represents a peptide of a  
XX CC cysteine protease epitope region relating to the antiallergic

CC compositions of the invention  
XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 61; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSCN 10  
||| ||| ||| ||| |||  
Db 1 RMQGGCGSCN 10

## RESULT 2

ABB98534  
ID ABB98534 standard; peptide; 10 AA.

XX AC ABB98534;

XX DT 13-DEC-2002 (first entry)

XX DE Cysteine protease epitope #1.

XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunotherapy; allergen; allergic hypersensitivity reaction;  
KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
KW cysteine protease.

XX OS Dermatophagoides pteronyssinus.

XX PN FR2822709-A1.

XX PD 04-OCT-2002.

XX PF 03-MAY-2001; 2001FR-00005929.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-735037/80.

XX Antiallergic composition, useful for preventing and treating e.g. asthma,  
PT rhinitis or eczema, containing at least two of allergen, antihistamine  
PT and histamine synthesis inhibitor.

XX PS Claim 8; Page 6; 33pp; French.

XX The present invention relates to an antiallergic pharmaceutical  
CC composition (I) comprising a pharmaceutical carrier containing an active  
CC agent combination of at least two of: an allergen; an antihistamine; and  
CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
CC rhinitis or allergic atopic eczema, in babies, children or adults. The  
CC present sequence is a peptide fragment (epitope) of cysteine protease  
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the  
CC invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 61; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSCN 10  
||| ||| ||| ||| |||  
Db 1 RMQGGCGSCN 10

## RESULT 3

AAR77350

ID AAR77350 standard; peptide; 8 AA.

XX AC AAR77350;

XX DT 14-MAY-1996 (first entry)

XX DE Cysteine proteinase derived peptide #1.

XX KW Primer; amplify; polymerase chain reaction; PCR; Haemonchus contortus;  
KW cysteine proteinase; DM.2; DM.3; DM.4; human; DM.4a; DM.5; antigen;  
KW vaccine; helminth; parasite; DM.1; ruminant; gastro-intestinal tract.

XX OS Haemonchus contortus.

XX PN WO9526402-A1.

XX PD 05-OCT-1995.

XX PF 24-MAR-1995; 95WO-GB000665.

XX PR 25-MAR-1994; 94GB-00005925.

XX PR 25-MAR-1994; 94GB-00005990.

XX (MLCW ) MALLINCKRODT VETERINARY INC.

XX PI Knox DP, Smith SK, Smith WD, Redmond D, Murray J;

XX WPI; 1995-351322/45.

XX DR N-PSDB; AAQ94240.

XX PT Protective helminth parasite antigen - used in vaccine directed against  
PT parasitic nematodes of mammalian gastro-intestinal tract e.g. Haemonchus  
PT contortus.

XX PS Example 16; Fig 15; 79pp; English.

XX The sequences given in AAR77350-53 are peptides derived from the  
CC canonical Haemonchus contortus cysteine proteinase molecule which were  
CC used in the design of the primers given in AAQ94240-43. These primers  
CC were used in the cloning of cDNA fragments from the cysteine proteinase  
CC gene, such as DM.1, DM.2, DM.3, DM.4, DM.4a and DM.5 (see also AAQ94246-  
CC 51). The amplified fragments may be expressed in a recombinant cell for  
CC the production of antigens. These antigens may be used in the preparation  
CC of a vaccine against helminth parasites in a human or non-human animal

XX SQ Sequence 8 AA;

Query Match 60.7%; Score 37; DB 2; Length 8;

Best Local Similarity 85.7%; Pred. No. 1.8e+06;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 QGCGGSC 9

Db 1 QGCGGSC 7

## RESULT 4

ID ADI11045

AD I ADI11045 standard; peptide; 9 AA.

XX AC ADI11045;

XX DT 15-APR-2004 (first entry)

XX DE Somatostatin analogue peptide SEQ ID NO:60.

XX KW dihydropyridine/pyridinium salt type redox moiety; retina; lipophilic;  
KW antidiabetic; ophthalmological; diabetic retinopathy;  
KW growth factor inhibitory activity; somatostatin; somatostatin analogue;  
KW octreotide; lanreotide; sequential metabolism;  
KW insulin dependent diabetes mellitus.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "Xaa = Nic"  
 FT Disulfide-bond 4. .7  
 FT Modified-site 9  
 FT /note= "Xaa = OH"  
 XX WO9912572-A1.  
 PN 18-MAR-1999.  
 XX PD  
 XX PF 01-SEP-1998; 98WO-US017987.  
 XX PR 10-SEP-1997; 97US-0058423P.  
 XX PA (UYFL ) UNIV FLORIDA.  
 XX PI Bodor NS, Grant MB;  
 XX WPI; 1999-263364/22.  
 XX DR  
 XX PT New peptide derivatives with redox targetor moiety are useful in the  
 PT treatment of diabetic retinopathy.  
 XX  
 PS Example 6; SEQ ID NO 60; 186pp; English.  
 XX  
 CC The present invention describes peptide derivatives (I) comprising a  
 CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides  
 CC to the retina, a bulky lipophilic function and an amino  
 CC acid/dipeptide/tripeptide spacer. Also described: (1) the preparation of  
 CC (I); and (2) intermediate quaternary salts (II). (I) have antidiabetic  
 CC and ophthalmological activities. (I) can be used in the treatment or  
 CC prevention of diabetic retinopathy by delivery of peptides with growth  
 CC factor inhibitory activity (e.g. somatostatin analogues such as  
 CC octreotide and lanreotide) to the retina by sequential metabolism. It is  
 CC envisaged that (I) will be useful in the treatment of insulin dependent  
 CC diabetes mellitus patients for critical periods in diabetic retinopathy  
 CC disease progression before laser photocoagulation is indicated,  
 CC preferably for 1-4 month intervals when a patient is experiencing severe  
 CC non-proliferative diabetic retinopathy or is found to have low risk  
 CC diabetic retinopathy. The present sequence represents a somatostatin  
 CC analogue peptide, which is used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 59.0%; Score 36; DB 2; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 |||||  
 Db 2 GCGGSC 7

RESULT 5  
 ADI11044  
 ID ADI11044 standard; peptide; 9 AA.  
 XX  
 AC ADI11044;

XX 15-APR-2004 (first entry)

XX Somatostatin analogue peptide SEQ ID NO:59.

XX dihydropyridine/pyridinium salt type redox moiety; retina; lipophilic;  
 KW antidiabetic; ophthalmological; diabetic retinopathy;  
 KW growth factor inhibitory activity; somatostatin; somatostatin analogue;  
 KW octreotide; lanreotide; sequential metabolism;  
 KW insulin dependent diabetes mellitus.  
 XX Synthetic.  
 OS

XX FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "Xaa = Trig"  
 FT Disulfide-bond 4. .7  
 FT Modified-site 9  
 FT /note= "Xaa = OH"  
 XX WO9912572-A1.  
 PN 18-MAR-1999.  
 XX PD  
 XX PF 01-SEP-1998; 98WO-US017987.  
 XX PR 10-SEP-1997; 97US-0058423P.  
 XX PA (UYFL ) UNIV FLORIDA.  
 XX PI Bodor NS, Grant MB;  
 XX WPI; 1999-263364/22.  
 XX DR  
 XX PT New peptide derivatives with redox targetor moiety are useful in the  
 PT treatment of diabetic retinopathy.  
 XX  
 PS Example 6; SEQ ID NO 59; 186pp; English.  
 XX  
 CC The present invention describes peptide derivatives (I) comprising a  
 CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides  
 CC to the retina, a bulky lipophilic function and an amino  
 CC acid/dipeptide/tripeptide spacer. Also described: (1) the preparation of  
 CC (I); and (2) intermediate quaternary salts (II). (I) have antidiabetic  
 CC and ophthalmological activities. (I) can be used in the treatment or  
 CC prevention of diabetic retinopathy by delivery of peptides with growth  
 CC factor inhibitory activity (e.g. somatostatin analogues such as  
 CC octreotide and lanreotide) to the retina by sequential metabolism. It is  
 CC envisaged that (I) will be useful in the treatment of insulin dependent  
 CC diabetes mellitus patients for critical periods in diabetic retinopathy  
 CC disease progression before laser photocoagulation is indicated,  
 CC preferably for 1-4 month intervals when a patient is experiencing severe  
 CC non-proliferative diabetic retinopathy or is found to have low risk  
 CC diabetic retinopathy. The present sequence represents a somatostatin  
 CC analogue peptide, which is used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 59.0%; Score 36; DB 2; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 |||||  
 Db 2 GCGGSC 7

RESULT 6  
 ABUS5751  
 ID ABUS5751 standard; peptide; 9 AA.  
 XX  
 AC ABUS5751;

XX 18-MAR-2003 (first entry)

XX Polylinker peptide #3 relating to invention of M. tuberculosis antigens.

XX Mycobacterium tuberculosis antigenic polypeptide; immune response;  
 KW tuberculosis infection; polylinker peptide.  
 XX Synthetic.  
 OS  
 XX US6465633-B1.  
 PN  
 XX

PI	Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX	
DR	WPI; 2003-897524/82.
XX	
PT	New fusion proteins of Mycobacterium tuberculosis for diagnosing, preventing or treating tuberculosis infection or in enhancing immune responses in M. tuberculosis.
XX	
PS	Disclosure; SEQ ID NO 43; 135pp; English.
XX	
CC	The present invention relates to fusion proteins of Mycobacterium tuberculosis antigens, and the polynucleotide sequences encoding them. The sequences of the invention are useful in a method for preventing tuberculosis by administering to a subject an amount of the fusion protein or the polynucleotide that encodes the fusion protein. Also disclosed is a pharmaceutical composition comprising the fusion protein or the polynucleotide sequence encoding it. The fusion protein induces an immune response to M. tuberculosis and can be used in the diagnosis, prevention, and treatment of tuberculosis infection. The present sequence represents a flexible polylinker peptide.
XX	
SQ	Sequence 9 AA;
QY	Query Match 59.0%; Score 36; DB 7; Length 9; Best Local Similarity 83.3%; Pred. No. 1.8e+06; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	4 GCGGSC 9 3 GCGGSC 8
RESULT 8	
AD111047	
ID	AD111047 standard; peptide; 10 AA.
XX	
AC	AD111047;
XX	
DT	15-APR-2004 (first entry)
XX	
DE	Somatostatin analogue peptide SEQ ID NO:62.
XX	
KW	dihydropyridine/pyridinium salt type redox moiety; retina; lipophilic; antidiabetic; ophthalmological; diabetic retinopathy;
KW	growth factor inhibitory activity; somatostatin; somatostatin analogue;
KW	octreotide; lanreotide; sequential metabolism;
KW	insulin dependent diabetes mellitus.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "Xaa = TFA"
FT	Modified-site 2 /note= "Xaa = Nic"
FT	Disulfide-bond 5..8
FT	Modified-site 10 /note= "Xaa = OH"
XX	
PN	WO9912572-A1.
XX	
PD	18-MAR-1999.
XX	
PF	01-SEP-1998; 98WO-US017987.
XX	
PR	10-SEP-1997; 97US-0058423P.
XX	
PA	(UYFL ) UNIV FLORIDA.
XX	
PI	Bodor NS, Grant MB;
XX	
DR	WPI; 1999-263364/22.
XX	



PT New peptide derivatives with redox targetor moiety are useful in the  
 PT treatment of diabetic retinopathy.

XX Example 10; SEQ ID NO 62; 186pp; English.

PS The present invention describes peptide derivatives (I) comprising a  
 CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides  
 CC to the retina, a bulky lipophilic function and an amino  
 CC acid/dipeptide/tripeptide spacer. Also described: (I) the preparation of  
 CC (1); and (2) intermediate quaternary salts (II). (I) have antidiabetic  
 CC and ophthalmological activities. (I) can be used in the treatment or  
 CC prevention of diabetic retinopathy by delivery of peptides with growth  
 CC factor inhibitory activity (e.g. somatostatin analogues such as  
 CC octreotide and lanreotide) to the retina by sequential metabolism. It is  
 CC envisaged that (I) will be useful in the treatment of insulin dependent  
 CC diabetes mellitus patients for critical periods in diabetic retinopathy  
 CC disease progression before laser photocoagulation is indicated.  
 CC preferably for 1-4 month intervals when a patient is experiencing severe  
 CC non-proliferative diabetic retinopathy or is found to have low risk  
 CC diabetic retinopathy. The present sequence represents a somatostatin  
 CC analogue peptide, which is used in the exemplification of the present  
 CC invention.

XX Sequence 10 AA;

Query Match 59.0%; Score 36; DB 2; Length 10;  
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGCGSC 9  
 Db 3 GGCGGC 8

RESULT 9

AAG86614  
 ID AAG86614 standard; peptide; 10 AA.

XX AAG86614;

DT 11-SEP-2001 (first entry)

XX Saccharomyces cerevisiae peptide, SEQ ID NO: 1563.

XX Saccharomyces cerevisiae; complementary peptide; peptide identification;  
 KW drug discovery; drug design.

XX Saccharomyces cerevisiae.

XX WO200142276-A1.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB004773.

XX 13-DEC-1999; 99GB-00029471.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis of protein and nucleotide  
 PT sequence databases, useful in drug design.

PS Example 3; Page 245; 488pp; English.

XX The invention relates to the identification of complementary peptides by  
 CC analysis of protein and nucleotide sequence databases from higher  
 CC eukaryotic genomes, excluding human and plants. The specific  
 CC complementary peptides interact with their relevant target proteins  
 CC encoded in the eukaryote genome. The peptides may be used as reagents and

CC drugs for drug discovery and as lead ligands for drug design and  
 CC development. The present sequence is a complementary peptide from  
 CC Saccharomyces cerevisiae

XX Sequence 10 AA;

Query Match 52.5%; Score 32; DB 4; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 MOGGCGS 8  
 Db 4 VSGGCGS 10

RESULT 10

AAW41566

ID AAW41566 standard; peptide; 8 AA.

XX AAW41566;

DT 27-APR-1998 (first entry)

XX Human calpain partial peptide.

XX Calpain; human; leukocyte; calcium dependent cysteine protease;  
 KW screening; activator; inhibitor; treatment; prevention; cancer;  
 KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;  
 KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;  
 KW ischaemic heart disease; atherosclerosis; arthritis.

XX Homo sapiens.

XX EP799892-A2.

XX 08-OCT-1997.

XX 03-APR-1997; 97BP-00105508.

XX 05-APR-1996; 96JP-00083649.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Shintani Y, Nishi K, Kawamoto T;

XX WPI; 1997-482674/45.

XX N-PSDB; AAV04204.

XX Human calpain protein and related DNA - useful for drug screening and  
 PT treating cancer, stroke, etc.

XX Disclosure; Page 29; 43pp; English.

XX The present sequence is a calpain partial peptide. Calpain is a human  
 CC leukocyte derived calcium dependent cysteine protease, which can be used  
 CC to screen for compounds that activate or inhibit its proteolytic  
 CC activity. Calpain DNA can be used to treat or prevent cancer, cerebral  
 CC apoplexy, cerebral infarction, subarachnoid haemorrhage, Alzheimer's  
 CC disease, myodystrophy, cataracts, ischaemic heart disease,  
 CC atherosclerosis, arthritis or collagen disease

XX Sequence 8 AA;

Query Match 50.8%; Score 31; DB 2; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGGCGSC 9

Db 2 QGGGDC 8

RESULT 11

AAB97506  
 ID AAB97506 standard; peptide; 8 AA.  
 AC AAB97506;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Peptide nucleic acid peptide fragment #4.  
 XX  
 KW Peptide nucleic acid; PNA; therapy; infection; cancer; restenosis;  
 KW asthma; autoimmune disorder; endocrinological disorder; renal failure;  
 KW neurological disease; acromegaly; sickle cell anaemia;  
 KW polyamide backbone.  
 XX  
 OS Synthetic.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT /note= "optionally bound to AAH23598"  
 FT Misc-difference 5  
 FT /label= OTHER  
 FT /note= "optionally D-form residue"  
 FT  
 XX US6180767-B1.  
 XX  
 XX 30-JAN-2001.  
 XX  
 PF 07-JAN-1997; 97US-00779072.  
 XX  
 PR 11-JAN-1996; 96US-0009747P.  
 XX  
 PA (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Wickstrom E, Basu S;  
 XX  
 DR WPI; 2001-342005/36.  
 XX  
 XX Novel conjugate, useful for killing pathogenic organisms and for  
 PT inhibiting gene expression, comprising a peptide nucleic acid oligomer  
 PT conjugated to ligand capable of binding to a cell surface receptor via a  
 PT chemical bond or linker.  
 XX  
 PS Example 1; Col 22; 25pp; English.  
 XX  
 CC The present invention provides a number of peptide nucleic acids (PNAs)  
 CC capable of binding to a cell surface receptor, where the oligomer and the  
 CC peptide are linked by a chemical bond or an amino acid linker. The  
 CC oligomer may have a polyamide, polythioamide, polysulfonamide or  
 CC polysulfonamide backbone. The PNAs of the invention can be used in  
 CC therapy, including the treatment of infections, cancer, autoimmune  
 CC diseases, renal failure, endocrinological disorders, acromegaly,  
 CC neurological diseases and sickle cell anaemia. The present sequence is an  
 CC example of a peptide for use in a PNA  
 XX  
 SQ Sequence 8 AA;  
 XX  
 Query Match 50.8%; Score 31; DB 4; Length 8;  
 Best Local Similarity 60.7%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GCGCSC 9  
 DB 3 GGCAC 8  
 ||| :  
 RESULT 12  
 ABB08544  
 ID ABB08544 standard; protein; 6 AA.  
 XX  
 AC ABB08544;  
 XX  
 DT 23-MAY-2002 (first entry)  
 XX

XX Human HCCa1 PCR primer P4.  
 XX  
 KW HCCa1; liver cancer; cytostatic; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN CN1322732-A.  
 XX  
 PD 21-NOV-2001.  
 XX  
 PF 08-MAY-2000; 2000CN-00115595.  
 XX  
 PR 08-MAY-2000; 2000CN-00115595.  
 XX  
 PA (SHAN-) SHANGHAI DONGFANG LIVER & GALLBLADDER SU.  
 XX  
 PI Wang H, Zeng J, Wu M;  
 XX  
 DR WPI; 2002-148617/20.  
 XX  
 PT New liver cancer up expressing gene for treating liver cancer and other  
 PT diseases.  
 XX  
 PS Example 5; Page 16 (disclosure); 31pp; Chinese.  
 XX  
 CC The present invention discloses new human HCCa1 protein, the  
 CC polynucleotides encoding the polypeptide and the recombinant process to  
 CC produce the polypeptide. The present invention also discloses the method  
 CC of applying the medicine composite of the polypeptide in treating liver  
 CC cancer and other diseases. The present invention also discloses the  
 CC preparation process of HCCa1 protein specific antibody and its  
 CC application in diagnosing and treating diseases. The present invention  
 CC also discloses the application of the polynucleotides encoding the new  
 CC HCCa1 protein. The present sequence represents a peptide sequence  
 CC relating to human HCCa1 protein  
 XX  
 SQ Sequence 6 AA;  
 XX  
 Query Match 49.2%; Score 30; DB 5; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 CGGCN 10  
 DB 1 CGGCN 5  
 ||| :  
 RESULT 13  
 ABUS5750  
 ID ABUS5750 standard; peptide; 6 AA.  
 XX  
 AC ABUS5750;  
 XX  
 DT 18-MAR-2003 (first entry)  
 XX  
 DE Polylinker peptide #2 relating to invention of M. tuberculosis antigens.  
 XX  
 KW Mycobacterium tuberculosis antigenic polypeptide; immune response;  
 KW tuberculosis infection; polylinker peptide.  
 XX  
 OS Synthetic.  
 XX  
 PN US6465633-B1.  
 XX  
 PD 15-OCT-2002.  
 XX  
 PF 23-DEC-1999; 99US-00470191.  
 XX  
 PR 24-DEC-1998; 98US-0113952P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX

PI Skeiky Y;  
 XX WPI; 2003-147072/14.  
 XX  
 XX Novel isolated mycobacterial polynucleotide, useful for treating,  
 XX preventing or diagnosing Mycobacterium tuberculosis infection, for  
 PT producing Mycobacterium tuberculosis secretory polypeptides and DNA  
 PT vaccines.  
 XX  
 XX Disclosure; Col 89; 48pp; English.  
 XX  
 XX The present invention relates to the isolation of polynucleotide  
 CC sequences encoding Mycobacterium tuberculosis antigenic polypeptides. The  
 CC polynucleotide sequences of the invention are useful for treating,  
 CC preventing, and diagnosing M. tuberculosis infection, for producing M.  
 CC tuberculosis secretory polypeptides, for producing DNA vaccines, for  
 CC diagnostic purposes, as molecular probes or primers to detect the  
 CC presence of bacteria in a biological sample, for inducing and/or  
 CC enhancing immune responses to M. tuberculosis, and in gene therapy.  
 CC ABU55749-ABU55754 represent flexible polylinker peptides. Note: The  
 CC present sequence is given in the sequence listing but is not mentioned  
 CC elsewhere in the specification  
 XX  
 XX Sequence 6 AA;  
 SQ  
 Query Match 49.2%; Score 30; DB 6; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 5 GCGGC 9  
 Db |||||  
 1 GCGGC 5

RESULT 14  
 ADF69792  
 ID ADF69792 standard; peptide; 6 AA.  
 XX  
 XX ADF69792;  
 AC  
 XX  
 XX 12-FEB-2004 (first entry)  
 DT  
 XX  
 XX M. tuberculosis fusion protein associated flexible polylinker peptide #1.  
 DE  
 XX  
 XX Fusion protein; Mycobacterium tuberculosis antigen;  
 KW tuberculosis infection; immune response; tuberculostatic;  
 KW flexible polylinker.  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX US2003147911-A1.  
 PN  
 XX  
 XX 07-AUG-2003.  
 PD  
 XX  
 XX 05-FEB-2003; 2003US-00359460.  
 PF  
 XX  
 XX 13-MAR-1997; 97US-00818112.  
 PR  
 XX 01-OCT-1997; 97US-00942578.  
 PR  
 XX 18-FEB-1998; 98US-00025197.  
 PR  
 XX 07-APR-1998; 98US-00056556.  
 PR  
 XX 30-DEC-1998; 98US-00223040.  
 PR  
 XX 07-APR-1999; 99US-00287849.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX  
 XX  
 XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
 PI  
 XX  
 XX WPI; 2003-897524/82.  
 DR  
 XX  
 XX New fusion proteins of Mycobacterium tuberculosis for diagnosing,  
 PT preventing or treating tuberculosis infection or in enhancing immune  
 PT responses in M. tuberculosis.  
 XX

PS Disclosure; SEQ ID NO 42; 135pp; English.  
 XX  
 XX The present invention relates to fusion proteins of Mycobacterium  
 CC tuberculosis antigens, and the polynucleotide sequences encoding them.  
 CC The sequences of the invention are useful in a method for preventing  
 CC tuberculosis by administering to a subject an amount of the fusion  
 CC protein or the polynucleotide that encodes the fusion protein. Also  
 CC disclosed is a pharmaceutical composition comprising the fusion protein  
 CC or the polynucleotide sequence encoding it. The fusion protein induces an  
 CC immune response to M. tuberculosis and can be used in the diagnosis,  
 CC prevention, and treatment of tuberculosis infection. The present sequence  
 CC represents a flexible polylinker peptide.  
 XX  
 XX Sequence 6 AA;  
 SQ  
 Query Match 49.2%; Score 30; DB 7; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 5 GCGSC 9  
 Db |||||  
 1 GCGSC 5

RESULT 15  
 AAB97503  
 ID AAB97503 standard; peptide; 8 AA.  
 XX  
 XX AAB97503;  
 AC  
 XX  
 XX 14-AUG-2001 (first entry)  
 DT  
 XX  
 XX Peptide nucleic acid peptide fragment #1.  
 DE  
 XX  
 XX Peptide nucleic acid; PNA; therapy; infection; cancer; restenosis;  
 KW asthma; autoimmune disorder; endocrinological disorder; renal failure;  
 KW neurological disease; acromegaly; sickle cell anaemia;  
 KW polyamide backbone.  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /label= OTHER  
 FT /note= "optionally bound to AAH23598"  
 FT Modified-site 5 /label= OTHER  
 FT /note= "optionally 4-methoxybenzyl-Cys, optionally D-form  
 FT residue"  
 FT Modified-site 6 /label= OTHER  
 FT /note= "optionally modified by benzyl, optionally D-form  
 FT residue"  
 FT Modified-site 7 /label= OTHER  
 FT /note= "optionally modified by phenylmethoxycarbonyl,  
 FT optionally D-form residue"  
 FT Modified-site 8 /label= OTHER  
 FT /note= "optionally modified by phenylmethoxycarbonyl and  
 FT resin, optionally D-form residue"  
 XX  
 XX US6180767-B1.  
 PN  
 XX  
 XX 30-JAN-2001.  
 PD  
 XX  
 XX 07-JAN-1997; 97US-00779072.  
 PF  
 XX  
 XX 11-JAN-1996; 96US-0009747P.  
 PR  
 XX  
 XX (UYJE-) UNIV JEFFERSON THOMAS.  
 PA  
 XX  
 XX Wickstrom E, Basu S;  
 PI

XX WPI; 2001-342005/36.  
DR  
XX  
PT Novel conjugate, useful for killing pathogenic organisms and for  
PT inhibiting gene expression, comprising a peptide nucleic acid oligomer  
PT conjugated to ligand capable of binding to a cell surface receptor via a  
PT chemical bond or linker.  
XX  
PS Claim 14; Col 19; 25pp; English.  
XX  
CC The present invention provides a number of peptide nucleic acids (PNAs)  
CC capable of binding to a cell surface receptor, where the oligomer and the  
CC peptide are linked by a chemical bond or an amino acid linker. The  
CC oligomer may have a polyamide, polychicamide, polysulfonamide or  
CC polysulfonamide backbone. The PNAs of the invention can be used in  
CC therapy, including the treatment of infections, cancer, autoimmune  
CC diseases, renal failure, endocrinological disorders, acromegaly,  
CC neurological diseases and sickle cell anaemia. The present sequence is an  
CC example of a peptide for use in a PNA  
XX  
SQ Sequence 8 AA;  
Query Match 49.2%; Score 30; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY . 4 GGCGSC 9  
Db |||||  
3 GGCSKC 8  
Search completed: May 19, 2005, 17:59:05  
Job time : 147 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:59:41 ; Search time 89.5 Seconds  
(without alignments)  
37.375 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMOGGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 185062

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	10	10	US-09-867-159A-3
2	36	59.0	7	15	US-10-175-833-59
3	36	59.0	7	15	US-10-175-833-60
4	36	59.0	7	15	US-10-175-833-62
5	36	59.0	9	9	US-09-287-849-43
6	36	59.0	9	14	US-10-359-460-43
7	36	59.0	9	15	US-10-359-459-7
8	33	54.1	8	14	US-10-163-415-2
9	31.5	51.6	8	17	US-10-678-961B-62
10	31.5	51.6	8	17	US-10-702-228A-62
11	30	49.2	6	9	US-09-287-849-42
12	30	49.2	6	14	US-10-359-460-42
13	30	49.2	6	15	US-10-359-459-6

14	30	49.2	10	9	US-09-765-086-57	Sequence 57, Appl
15	30	49.2	10	9	US-09-779-308-222	Sequence 222, App
16	30	49.2	10	9	US-09-779-308-324	Sequence 324, App
17	30	49.2	10	9	US-09-779-308-436	Sequence 436, App
18	30	49.2	10	9	US-09-779-308-546	Sequence 546, App
19	30	49.2	10	9	US-09-779-308-640	Sequence 640, App
20	30	49.2	10	14	US-10-264-374-57	Sequence 57, Appl
21	30	49.2	10	14	US-10-375-992-57	Sequence 57, Appl
22	30	49.2	10	15	US-10-264-374-57	Sequence 57, Appl
23	30	49.2	10	16	US-10-375-992-57	Sequence 57, Appl
24	30	49.2	10	17	US-10-838-289-593	Sequence 593, App
25	29	47.5	8	13	US-10-010-408-4	Sequence 4, Appli
26	29	47.5	8	15	US-10-149-138-2467	Sequence 2467, Ap
27	29	47.5	8	15	US-10-149-138-3195	Sequence 3195, Ap
28	29	47.5	8	15	US-10-311-129-26	Sequence 26, Appl
29	29	47.5	8	16	US-10-149-138-2467	Sequence 2467, Ap
30	29	47.5	8	16	US-10-149-138-3195	Sequence 3195, Ap
31	29	47.5	9	15	US-10-149-138-855	Sequence 855, App
32	29	47.5	9	15	US-10-149-138-2495	Sequence 2495, Ap
33	29	47.5	9	15	US-10-149-138-3217	Sequence 3217, Ap
34	29	47.5	9	16	US-10-149-138-855	Sequence 855, App
35	29	47.5	9	16	US-10-149-138-2495	Sequence 2495, Ap
36	29	47.5	9	16	US-10-149-138-3217	Sequence 3217, Ap
37	29	47.5	9	16	US-10-149-138-4122	Sequence 4122, Ap
38	29	47.5	10	9	US-09-765-086-52	Sequence 52, Appl
39	29	47.5	10	10	US-09-572-404B-2757	Sequence 2757, Ap
40	29	47.5	10	10	US-09-572-404B-2759	Sequence 2759, Ap
41	29	47.5	10	10	US-09-572-404B-2761	Sequence 2761, Ap
42	29	47.5	10	10	US-09-572-404B-3780	Sequence 3780, Ap
43	29	47.5	10	14	US-10-264-374-52	Sequence 52, Appl
44	29	47.5	10	14	US-10-375-992-52	Sequence 52, Appl
45	29	47.5	10	14	US-10-375-992-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1

US-09-867-159A-3  
; Sequence 3, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE OF INVENTION: and at least one anti-histamine compound  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Comprises epitope from cystine protease.

US-09-867-159A-3

Query Match 100.0%; Score 61; DB 10; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.03; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0;

Qy 1 RMOGGCGSCN 10  
Db 1 RMOGGCGSCN 10

```
RESULT 2
; US-10-175-833-59
; Sequence 59, Application US/10175833
; Publication No. US20030211981A1
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas Stephen
; APPLICANT: BARTOLOMEO, Maria
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT
; TITLE OF INVENTION: OF DIABETIC RETINOPATHY
; FILE REFERENCE: 028724-109
; CURRENT APPLICATION NUMBER: US/10/175,833
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/144,991
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/058,423
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)
; OTHER INFORMATION: Amino acid 1 is attached by Trig.
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (7)
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.
; NAME/KEY: DISULFID
; LOCATION: (3)..(6)
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are
; OTHER INFORMATION: attached by a non-peptidal disulfide bond.
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative
; US-10-175-833-59

Query Match          59.0%; Score 36; DB 15; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.3e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 1 GCGGSC 6

RESULT 3
; US-10-175-833-60
; Sequence 60, Application US/10175833
; Publication No. US20030211981A1
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas Stephen
; APPLICANT: BARTOLOMEO, Maria
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT
; TITLE OF INVENTION: OF DIABETIC RETINOPATHY
; FILE REFERENCE: 028724-109
; CURRENT APPLICATION NUMBER: US/10/175,833
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/144,991
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/058,423
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)
; OTHER INFORMATION: Amino acid 1 is attached by Nic.
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (7)
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.
; NAME/KEY: DISULFID
; LOCATION: (3)..(6)
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are
; OTHER INFORMATION: attached by a non-peptidal disulfide bond.
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative
; US-10-175-833-60

Query Match          59.0%; Score 36; DB 15; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.3e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 1 GCGGSC 6

RESULT 4
; US-10-175-833-62
; Sequence 62, Application US/10175833
; Publication No. US20030211981A1
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas Stephen
; APPLICANT: BARTOLOMEO, Maria
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT
; TITLE OF INVENTION: OF DIABETIC RETINOPATHY
; FILE REFERENCE: 028724-109
; CURRENT APPLICATION NUMBER: US/10/175,833
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/144,991
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/058,423
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)
; OTHER INFORMATION: Amino acid 1 is attached by Nic.
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (7)
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.
; NAME/KEY: DISULFID
; LOCATION: (3)..(6)
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are
; OTHER INFORMATION: attached by a non-peptidal disulfide bond.
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative
; US-10-175-833-62

Query Match          59.0%; Score 36; DB 15; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.3e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 1 GCGGSC 6

RESULT 5
; US-09-287-849-43
; Sequence 43, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
```

```
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (7)
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.
; NAME/KEY: DISULFID
; LOCATION: (3)..(6)
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are
; OTHER INFORMATION: attached by a non-peptidal disulfide bond.
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative
; US-10-175-833-60
```

```
Query Match          59.0%; Score 36; DB 15; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.3e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 1 GCGGSC 6
```

```
RESULT 4
; US-10-175-833-62
; Sequence 62, Application US/10175833
; Publication No. US20030211981A1
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas Stephen
; APPLICANT: BARTOLOMEO, Maria
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT
; TITLE OF INVENTION: OF DIABETIC RETINOPATHY
; FILE REFERENCE: 028724-109
; CURRENT APPLICATION NUMBER: US/10/175,833
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/144,991
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: US 60/058,423
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (1)
; OTHER INFORMATION: Amino acid 1 is attached by Nic.
; FEATURE:
; NAME/KEY: BINDING
; LOCATION: (7)
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.
; NAME/KEY: DISULFID
; LOCATION: (3)..(6)
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are
; OTHER INFORMATION: attached by a non-peptidal disulfide bond.
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative
; US-10-175-833-62
```

```
Query Match          59.0%; Score 36; DB 15; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.3e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 1 GCGGSC 6
```

```
RESULT 5
; US-09-287-849-43
; Sequence 43, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
```

<pre>; APPLICANT: Reed, Steven G. ; APPLICANT: Skeiky, Yasir A.W. ; APPLICANT: Dillon, Davin C. ; APPLICANT: Alderson, Mark ; APPLICANT: Campos-Neto, Antonio ; APPLICANT: Corixa Corporation ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens ; TITLE OF INVENTION: and Their Uses ; FILE REFERENCE: 014058-009020US ; CURRENT APPLICATION NUMBER: US 09/287,849 ; CURRENT FILING DATE: 1999-04-07 ; PRIOR APPLICATION NUMBER: US 08/818,112 ; PRIOR FILING DATE: 1997-03-13 ; PRIOR APPLICATION NUMBER: US 08/942,578 ; PRIOR FILING DATE: 1997-10-01 ; PRIOR APPLICATION NUMBER: US 09/025,197 ; PRIOR FILING DATE: 1998-02-18 ; PRIOR APPLICATION NUMBER: US 09/056,556 ; PRIOR FILING DATE: 1998-04-07 ; PRIOR APPLICATION NUMBER: US 09/223,040 ; PRIOR FILING DATE: 1998-12-30 ; NUMBER OF SEQ ID NOS: 46 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 43 ; LENGTH: 9 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence:flexible ; OTHER INFORMATION: polylinker US-09-287-849-43</pre>	<pre>Query Match          59.0%; Score 36; DB 9; Length 9; Best Local Similarity 83.3%; Pred. No. 1.3e+06; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0  Qy      4 GCGGSC 9               Db      3 GCGGGC 8  RESULT 6 US-10-359-460-43 ; Sequence 43, Application US/10359460 ; Publication No. US2003014791A1 ; GENERAL INFORMATION: ; APPLICANT: Reed, Steven G. ; APPLICANT: Skeiky, Yasir A.W. ; APPLICANT: Dillon, Davin C. ; APPLICANT: Alderson, Mark ; APPLICANT: Campos-Neto, Antonio ; APPLICANT: Corixa Corporation ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens ; TITLE OF INVENTION: and Their Uses ; FILE REFERENCE: 014058-009020US ; CURRENT APPLICATION NUMBER: US/10/359,460 ; CURRENT FILING DATE: 2003-02-05 ; PRIOR APPLICATION NUMBER: US/09/287,849 ; PRIOR FILING DATE: 1999-04-07 ; PRIOR APPLICATION NUMBER: US 08/818,112 ; PRIOR FILING DATE: 1997-03-13 ; PRIOR APPLICATION NUMBER: US 08/942,578 ; PRIOR FILING DATE: 1997-10-01 ; PRIOR APPLICATION NUMBER: US 09/025,197 ; PRIOR FILING DATE: 1998-02-18 ; PRIOR APPLICATION NUMBER: US 09/056,556 ; PRIOR FILING DATE: 1998-04-07 ; PRIOR APPLICATION NUMBER: US 09/223,040 ; PRIOR FILING DATE: 1998-12-30 ; NUMBER OF SEQ ID NOS: 46 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 43 ; LENGTH: 9 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence:flexible ; OTHER INFORMATION: polylinker US-10-359-460-43</pre>
<pre>; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence:flexible ; OTHER INFORMATION: polylinker US-10-359-460-43  Query Match          59.0%; Score 36; DB 14; Length 9; Best Local Similarity 83.3%; Pred. No. 1.3e+06; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0  Qy      4 GCGGSC 9               Db      3 GCGGGC 8  RESULT 7 US-10-359-459-7 ; Sequence 7, Application US/10359459 ; Publication No. US20040013677A1 ; GENERAL INFORMATION: ; APPLICANT: Skeiky, Yasir ; APPLICANT: Alderson, Mark ; APPLICANT: Campos-Neto, Antonio ; APPLICANT: Corixa Corporation ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens ; TITLE OF INVENTION: and Their Uses ; FILE REFERENCE: 014058-009010US ; CURRENT APPLICATION NUMBER: US/10/359,459 ; CURRENT FILING DATE: 2003-02-05 ; PRIOR APPLICATION NUMBER: US/09/223,040 ; PRIOR FILING DATE: 1998-12-30 ; NUMBER OF SEQ ID NOS: 10 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 7 ; LENGTH: 9 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence:flexible ; OTHER INFORMATION: polylinker US-10-359-459-7</pre>	<pre>Query Match          59.0%; Score 36; DB 15; Length 9; Best Local Similarity 83.3%; Pred. No. 1.3e+06; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0  Qy      4 GCGGSC 9               Db      3 GCGGGC 8  RESULT 8 US-10-163-415-2 ; Sequence 2, Application US/10163415 ; Publication No. US20030129204A1 ; GENERAL INFORMATION: ; APPLICANT: KNOX, DAVID PATRICK ; APPLICANT: SMITH, STUART KEVIN ; APPLICANT: SMITH, WILLIAM DAVID ; APPLICANT: REDMOND, DIANE ; APPLICANT: MURRAY, JACQUELINE ; TITLE OF INVENTION: VACCINES AGAINST HELMINTHIC PARASITES ; FILE REFERENCE: 1181-264 ; CURRENT APPLICATION NUMBER: US/10/163,415 ; CURRENT FILING DATE: 2002-06-07 ; PRIOR APPLICATION NUMBER: US 08/716418 ; PRIOR FILING DATE: 1996-09-20 ; PRIOR APPLICATION NUMBER: PCT/GB95/00665 ; PRIOR FILING DATE: 1995-03-24 ; PRIOR APPLICATION NUMBER: GB 9405925.0 ; PRIOR FILING DATE: 1994-03-25 ; PRIOR APPLICATION NUMBER: GB 9405990.4 ; LENGTH: 9</pre>

; PRIOR FILING DATE: 1994-03-25  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; NAME/KEY: misc feature  
; LOCATION: (6)-(6)  
; OTHER INFORMATION: The 'Xaa' at location 6 stands for Ser.  
; FEATURE:  
; OTHER INFORMATION: PCR primer: 508G  
US-10-163-415-2

Query Match 54.1%; Score 33; DB 14; Length 8;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGCGGSC 9  
|||  
Db 1 QGCGGXC 7

## RESULT 9

US-10-678-961B-62  
; Sequence 62, Application US/10678961B  
; Publication No. US20050074863A1  
; GENERAL INFORMATION:  
; APPLICANT: Slater, Michael R.  
; APPLICANT: Strauss, Ethan Edward  
; APPLICANT: Wood, Keith V.  
; APPLICANT: Hartnett, James Robert  
; APPLICANT: Promega Corporation  
; TITLE OF INVENTION: Vectors for Directional Cloning  
; FILE REFERENCE: 341.023US1  
; CURRENT APPLICATION NUMBER: US/10/678,961B  
; CURRENT FILING DATE: 2003-10-03  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: A synthetic peptide  
US-10-678-961B-62

Query Match 51.6%; Score 31.5; DB 17; Length 8;  
Best Local Similarity 85.7%; Pred. No. 1.3e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4 GCGGSCN 10  
|||  
Db 2 GCGG-CN 7

## RESULT 10

US-10-702-228A-62  
; Sequence 62, Application US/10702228A  
; Publication No. US20050074785A1  
; GENERAL INFORMATION:  
; APPLICANT: Slater, Michael R.  
; APPLICANT: Wood, Keith V.  
; APPLICANT: Hartnett, James Robert  
; APPLICANT: Promega Corporation  
; TITLE OF INVENTION: Vectors for Directional Cloning  
; FILE REFERENCE: 341.030US1  
; CURRENT APPLICATION NUMBER: US/10/702,228A  
; CURRENT FILING DATE: 2003-11-05  
; PRIOR APPLICATION NUMBER: 10/678,961  
; PRIOR FILING DATE: 2003-10-03  
; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: A synthetic peptide  
US-10-702-228A-62

Query Match 51.6%; Score 31.5; DB 17; Length 8;  
Best Local Similarity 85.7%; Pred. No. 1.3e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4 GCGGSCN 10  
|||  
Db 2 GCGG-CN 7

## RESULT 11

US-09-287-849-42  
; Sequence 42, Application US/09287849  
; Patent No. US20020009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:flexible  
; OTHER INFORMATION: polylinker  
US-09-287-849-42

Query Match 49.2%; Score 30; DB 9; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9  
|||  
Db 1 GCGGC 5

## RESULT 12

US-10-359-460-42  
; Sequence 42, Application US/10359460  
; Publication No. US20030147911A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark



; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/10/359,460  
; CURRENT FILING DATE: 2003-02-05  
; PRIOR APPLICATION NUMBER: US/09/287,849  
; PRIOR FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: flexible  
; OTHER INFORMATION: polylinker  
US-10-359-460-42

Query Match 49.2%; Score 30; DB 14; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9  
|||  
Db 1 GCGGC 5

RESULT 13  
US-10-359-459-6  
; Sequence 6, Application US/10359459  
; Publication No. US20040013677A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; FILE REFERENCE: 014058-009010US  
; CURRENT APPLICATION NUMBER: US/10/359,459  
; CURRENT FILING DATE: 2003-02-05  
; PRIOR APPLICATION NUMBER: US/09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: flexible  
; OTHER INFORMATION: polylinker  
US-10-359-459-6

Query Match 49.2%; Score 30; DB 15; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9  
|||  
Db 1 GCGGC 5

RESULT 14  
US-09-765-086-57  
; Sequence 57, Application US/09765086  
; Patent No. US20010046498A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Wadih, Arap  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Ellerbe, H. Michael  
; TITLE OF INVENTION: Chimeric Prostate-Homing Peptides With  
; TITLE OF INVENTION: Pro-Apoptotic Activity  
; FILE REFERENCE: P-LJ 3844  
; CURRENT APPLICATION NUMBER: US/09/765,086  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US 09/489,582  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 235  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 57  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-765-086-57

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 80.0%; Pred. No. 7.1e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CGSCN 10  
|||  
Db 1 CGECN 5

RESULT 15  
US-09-779-308-222  
; Sequence 222, Application US/09779308  
; Patent No. US20020150972A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 34PD7: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.4USU1  
; CURRENT APPLICATION NUMBER: US/09/779,308  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181,020  
; PRIOR FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 718  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-779-308-222

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 71.4%; Pred. No. 7.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMOGCG 7  
|||  
Db 1 RLOGGAG 7

Search completed: May 19, 2005, 18:19:58  
Job time : 90.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:53:34 ; Search time 29.5 Seconds  
(without alignments)  
25.305 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 115750

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	59.0	7	4	US-09-144-991B-59
2	36	59.0	7	4	US-09-144-991B-60
3	36	59.0	7	4	US-09-144-991B-62
4	36	59.0	9	4	US-09-470-191-94
5	36	59.0	9	4	US-09-223-040-7
6	36	59.0	9	4	US-09-287-849-43
7	31	50.8	8	2	US-08-835-099A-3
8	31	50.8	8	3	US-09-157-349-3
9	31	50.8	8	3	US-08-779-072A-6
10	30	49.2	6	4	US-09-470-191-93
11	30	49.2	6	4	US-09-223-040-6
12	30	49.2	6	4	US-09-287-849-42
13	30	49.2	8	3	US-08-779-072A-1
14	30	49.2	10	2	US-08-361-864-36
15	30	49.2	10	2	US-08-902-367-7
16	30	49.2	10	3	US-08-535-170-9
17	30	49.2	10	3	US-09-139-802-57
18	30	49.2	10	4	US-09-659-786-57
19	30	49.2	10	4	US-08-926-914-57
20	29	47.5	9	1	US-08-482-880-8
21	29	47.5	9	2	US-08-273-274-8
22	29	47.5	9	2	US-08-475-041-8
23	29	47.5	9	2	US-08-484-773-8
24	29	47.5	10	2	US-08-335-832-8
25	29	47.5	10	3	US-09-141-127-2
26	29	47.5	10	3	US-09-139-802-52
27	29	47.5	10	4	US-09-659-786-52

28	29	47.5	10	4	US-08-926-914-52	Sequence 52, Appl
29	28	45.9	5	1	US-08-467-607-10	Sequence 10, Appl
30	28	45.9	5	2	US-08-469-362-10	Sequence 10, Appl
31	28	45.9	5	2	US-08-850-392-10	Sequence 10, Appl
32	28	45.9	6	4	US-09-982-704-9	Sequence 9, Appl
33	28	45.9	7	3	US-08-827-171B-13	Sequence 13, Appl
34	28	45.9	7	4	US-09-588-995A-111	Sequence 111, App
35	28	45.9	7	4	US-09-598-062-13	Sequence 13, Appl
36	28	45.9	8	1	US-08-526-710-28	Sequence 28, Appl
37	28	45.9	8	3	US-08-862-855-28	Sequence 28, Appl
38	28	45.9	8	3	US-09-226-985-28	Sequence 28, Appl
39	28	45.9	8	3	US-09-227-906-28	Sequence 28, Appl
40	28	45.9	8	4	US-09-228-866-28	Sequence 28, Appl
41	28	45.9	9	3	US-08-997-802-10	Sequence 10, Appl
42	28	45.9	9	3	US-08-997-802-11	Sequence 11, Appl
43	28	45.9	10	3	US-09-139-802-32	Sequence 32, Appl
44	28	45.9	10	4	US-09-659-786-32	Sequence 32, Appl
45	28	45.9	10	4	US-08-926-914-32	Sequence 32, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-144-991B-59  
; Sequence 59, Application US/09144991B  
; Patent No. 6440933  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas Stephen  
; APPLICANT: BARTOLOMEO, Maria  
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
; FILE REFERENCE: 028724-109  
; CURRENT APPLICATION NUMBER: US/09/144,991B  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: US 60/058,423  
; PRIOR FILING DATE: 1997-09-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 59  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: BINDING  
; LOCATION: (1)  
; OTHER INFORMATION: Amino acid 1 is attached by Trig.  
; NAME/KEY: BINDING  
; LOCATION: (7)  
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
; NAME/KEY: DISULFID  
; LOCATION: (3)..(6)  
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
; OTHER INFORMATION: attached by a non-peptidic disulfide bond.  
; OTHER INFORMATION: Description of Unknown Organism: peptide derivative  
US-09-144-991B-59

Query Match 59.0%; Score 36; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
|||  
Db 1 GCGGCG 6

##### RESULT 2

US-09-144-991B-60  
; Sequence 60, Application US/09144991B  
; Patent No. 6440933  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas Stephen  
; APPLICANT: BARTOLOMEO, Maria

; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
; FILE REFERENCE: 028724-109  
; CURRENT APPLICATION NUMBER: US/09/144,991B  
; CURRENT FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: US 60/058,423  
; PRIOR FILING DATE: 1997-09-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 60  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: BINDING  
; LOCATION: (1)  
; OTHER INFORMATION: Amino acid 1 is attached by Nic.  
; NAME/KEY: BINDING  
; LOCATION: (7)  
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
; NAME/KEY: DISULFIDE  
; LOCATION: (3)..(6)  
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
; OTHER INFORMATION: attached by a non-peptidic disulfide bond.  
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
US-09-144-991B-60

Query Match 59.0%; Score 36; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
Db 1 GCGGCG 6

RESULT 3  
US-09-144-991B-62  
; Sequence 62, Application US/09144991B  
; Patent No. 6440933  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas Stephen  
; APPLICANT: BARTOLOMEO, Maria  
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
; FILE REFERENCE: 028724-109  
; CURRENT APPLICATION NUMBER: US/09/144,991B  
; CURRENT FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: US 60/058,423  
; PRIOR FILING DATE: 1997-09-10  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 62  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: BINDING  
; LOCATION: (1)  
; OTHER INFORMATION: Amino acid 1 is attached by Nic.  
; NAME/KEY: BINDING  
; LOCATION: (7)  
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
; NAME/KEY: DISULFID  
; LOCATION: (3)..(6)  
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
; OTHER INFORMATION: attached by a non-peptidic disulfide bond.  
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
US-09-144-991B-62

Query Match 59.0%; Score 36; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
Db 1 GCGGCG 6

RESULT 4  
US-09-470-191-94  
; Sequence 94, Application US/09470191  
; Patent No. 6465633  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods of Their Use in  
; TITLE OF INVENTION: the Treatment, Prevention and Diagnosis of Tuberculosis  
; FILE REFERENCE: 014058-008910US  
; CURRENT APPLICATION NUMBER: US/09/470,191  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: US 60/113,952  
; PRIOR FILING DATE: 1998-12-24  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 94  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: flexible polylinker  
US-09-470-191-94

Query Match 59.0%; Score 36; DB 4; Length 9;  
Best Local Similarity 83.3%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
Db 3 GCGGCG 8

RESULT 5  
US-09-223-040-7  
; Sequence 7, Application US/09223040  
; Patent No. 6544522  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009010US  
; CURRENT APPLICATION NUMBER: US/09/223,040  
; CURRENT FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:flexible  
; OTHER INFORMATION: polylinker  
US-09-223-040-7

Query Match 59.0%; Score 36; DB 4; Length 9;  
Best Local Similarity 83.3%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
Db 3 GCGGCG 8

```
RESULT 6
US-09-287-849-43
; Sequence 43, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible
; OTHER INFORMATION: polylinker
US-09-287-849-43

Query Match 59.0%; Score 36; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 3 GCGGCGC 8

RESULT 7
US-08-835-099A-3
; Sequence 3, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazunori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-835-099A-3

Query Match 50.8%; Score 31; DB 2; Length 8;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGCGGSC 9
Db 2 QGGLGDC 8

RESULT 8
US-09-157-349-3
; Sequence 3, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazunori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-835-099A-3
```

```
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-157-349-3

Query Match          50.8%; Score 31; DB 3; Length 8;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGCGGSC 9
Db 2 QGGGDC 8

RESULT 9
US-08-779-072A-6
; Sequence 6, Application US/08779072A
; Patent No. 6180767
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric
; APPLICANT: Basu, Soumitra
; TITLE OF INVENTION: PEPTIDE NUCLEIC ACID CONJUGATES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,072A
; FILING DATE: January 7, 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,747
; FILING DATE: January 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 6180767e
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-779-072A-6

Query Match          50.8%; Score 31; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9
Db 3 GGCAAC 8

RESULT 10
US-09-470-191-93
; Sequence 93, Application US/09470191
; Patent No. 6465633

; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods of Their Use in
; FILE REFERENCE: 014058-008910US
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 60/113,952
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: flexible polylinker
; US-09-470-191-93

Query Match          49.2%; Score 30; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9
Db 1 GCGGC 5

RESULT 11
US-09-223-040-6
; Sequence 6, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
; US-09-223-040-6

Query Match          49.2%; Score 30; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9
Db 1 GCGGC 5

RESULT 12
US-09-287-849-42
; Sequence 42, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
```

; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; FILE REFERENCE: and Their Uses  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:flexible  
; OTHER INFORMATION: polylinker  
US-09-287-849-42

Query Match 49.2%; Score 30; DB 4; Length 6;  
Best Local Similarity 80.0%; Pred. No. 4.le+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9  
Db 1 GCGGC 5

RESULT 13  
US-08-779-072A-1  
; Sequence 1, Application US/08779072A  
; Patent No. 6180767  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric  
; APPLICANT: Basu, Sumittra  
; TITLE OF INVENTION: PEPTIDE NUCLEIC ACID CONJUGATES  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.  
; STREET: Suite 1800, Two Penn Center Plaza  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,072A  
; FILING DATE: January 7, 1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,747  
; FILING DATE: January 11, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-14  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 6180767e  
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-779-072A-1  
Query Match 49.2%; Score 30; DB 3; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4.le+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 GCGSC 9  
Db 3 GCGSC 8  
RESULT 14  
US-08-361-864-36  
; Sequence 36, Application US/08361864  
; Patent No. 5977064  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T  
; APPLICANT: Lister-James, John  
; TITLE OF INVENTION: Multimeric Polyvalent Antithrombotic  
; TITLE OF INVENTION: Agents  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: 10 South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/361,864  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/955,466A  
; FILING DATE: 19921002  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5977064han, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 92,668  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 910-221-5317  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..5  
; OTHER INFORMATION: /label= Cyclic  
; OTHER INFORMATION: /note= "The sidechain sulfur of the 1st cysteine  
; OTHER INFORMATION: is protected by an -CH2CO- group, that also forms  
; OTHER INFORMATION: an amide bond with the N-terminus; the Y is the D  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 6..8  
; OTHER INFORMATION: /label= Tc-99m-binding  
; OTHER INFORMATION: /note= "The sidechain sulfur atom of each cysteine  
; OTHER INFORMATION: is protected by an acetamido group; the C-terminal  
; OTHER INFORMATION: cysteine is an amide"

US-08-361-864-36

Query Match 49.2%; Score 30; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 3 QGCGSC 9  
: ||| |  
Db 2 RGDCGC 8

RESULT 15

US-08-902-367-7  
; Sequence 7, Application US/08902367  
; Patent No. 597845  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T.  
; APPLICANT: Lister-James, John  
; APPLICANT: Civitello, Edgar R.  
; TITLE OF INVENTION: Radiolabeled Compounds for Thrombus  
; TITLE OF INVENTION: Imaging  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive Seventh Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/902,367  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/462,668  
; FILING DATE: 05-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5997845nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 90,1104-W  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312 913 0001  
; TELEFAX: 312 913 0002  
; TELEX:

INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..3  
; OTHER INFORMATION: /label= D-Tyr  
; OTHER INFORMATION: /note= "The tyrosine residue is in the D-stereo-  
; OTHER INFORMATION: chemical configuration"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..5  
; OTHER INFORMATION: /label= Cyclic  
; OTHER INFORMATION: /note= "The sidechain sulfur of the Cys  
; OTHER INFORMATION: residue is covalently linked to the amino  
; OTHER INFORMATION: terminus by a -CH2CO- group."  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 8..10  
; OTHER INFORMATION: /label= Tc-99m-chelator  
; OTHER INFORMATION: /note= "The sidechain sulfur atoms of both Cys

; OTHER INFORMATION: residues are each protected with an  
; OTHER INFORMATION: acetamidomethyl group"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 10  
; OTHER INFORMATION: /label= Amide  
; OTHER INFORMATION: /note= "The carboxyl terminus is modified to an  
; OTHER INFORMATION: amide"  
; US-08-902-367-7

Query Match 49.2%; Score 30; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 3 QGCGSC 9  
: ||| |  
Db 2 RGDCGC 8

Search completed: May 19, 2005, 18:11:52  
Job time : 31.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:50:19 ; Search time 12 Seconds  
(without alignments)  
80.181 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	32.7	10	2 S65387	cytochrome-c oxida
2	17	30.9	8	2 JS0316	leucokinin VI - Ma
3	16	29.1	7	1 NYPG7	hypothalamic hepta
4	16	29.1	9	2 PT0285	Ig heavy chain CRD
5	16	29.1	9	2 S55696	phosphoenolpyruvat
6	16	29.1	10	2 A43405	6-phosphofructo-2-
7	16	29.1	10	2 S33844	alpha-2-macroglobu
8	16	29.1	10	2 PT0291	Ig heavy chain CRD
9	16	29.1	10	2 A59272	peptide-N4-(N-acet
10	16	29.1	10	2 S23307	neurokinin A - rat
11	16	29.1	10	2 S23186	neurokinin A - Atl
12	15	27.3	7	2 A38081	amine oxidase (cop
13	15	27.3	7	2 A15398	choline oxidase (E
14	15	27.3	8	2 A44960	neuropeptide Led-C
15	15	27.3	8	2 B44960	neuropeptide Led-C
16	15	27.3	8	2 S08995	hypertrehalosemic
17	15	27.3	8	2 S08996	hypertrehalosemic
18	15	27.3	8	2 A49823	adipokinetic hormo
19	15	27.3	8	2 B49823	adipokinetic hormo
20	15	27.3	8	2 S15422	adipokinetic hormo
21	15	27.3	8	2 A43976	hypertrehalosemic
22	15	27.3	8	2 A43976	hypertrehalosemic
23	15	27.3	8	2 A58641	adipokinetic hormo
24	15	27.3	8	2 A05169	neuropeptide M-I -
25	15	27.3	9	2 S70345	amine oxidase (cop
26	15	27.3	9	2 PT0240	Ig heavy chain CRD
27	15	27.3	9	2 S13889	phosphoenolpyruvat
28	15	27.3	9	2 S77984	cytochrome-c oxida
29	15	27.3	10	1 ECLQ4M	tachykinin IV - mi

30 15 27.3 10 2 JC1416 hypertrehalosemic  
31 15 27.3 10 2 S09138 hypertrehalosemic  
32 15 27.3 10 2 PT0213 T-cell receptor al  
33 15 27.3 10 2 JQ0943 hypothetical 1.3K  
34 14 25.5 6 2 PQ0008 angiotensin-conver  
35 14 25.5 7 2 S71867 glutathione transf  
36 14 25.5 8 2 A59028 MHC class I histoc  
37 14 25.5 9 2 A61230 calsequestrin, car  
38 14 25.5 9 2 C41170 photosystem II pro  
39 14 25.5 9 2 PT0288 Ig heavy chain CRD  
40 14 25.5 9 2 PT0324 Ig heavy chain CRD  
41 14 25.5 9 2 D41978 callipressamide 4 -  
42 14 25.5 10 2 S10926 inhibit beta-A cha  
43 14 25.5 10 2 A60647 neuromedin C - bov  
44 13 23.6 6 2 A44916 mosquitoicidal toxi  
45 13 23.6 7 2 S21230 dermorphin (Trp-4,

#### ALIGNMENTS

##### RESULT 1

S65387

cytochrome-c oxidase (EC 1.9.3.1) chain VII b, cardiac - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 09-Jul-2004

C;Accession: S65387; S65386

R;Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.

Eur. J. Biochem. 230, 235-241, 1995

A;Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term

A;Reference number: S65372; MUID:95324529; PMID:7601105

A;Accession: S65387

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <SCH>

A;Cross-references: UNIPROT:P80431

A;Accession: S65386

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <SC2>

C;Keywords: cardiac muscle; heart; oxidoreductase

Query Match 32.7%; Score 18; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 1.6e+03;

Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYH 5

DB 7 PTFH 10

##### RESULT 2

JS0316

leucokinin VI - Madeira cockroach

C;Species: Leucophaea madeira (Madeira cockroach)

C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C;Accession: JS0316

R;Holman, G.M.; Cook, B.J.; Nachman, R.J.

Comp. Biochem. Physiol. C 88, 27-30, 1987

A;Title: Isolation, primary structure, and synthesis of leucokinsins V and VI: myotropic

A;Reference number: JS0315

A;Accession: JS0316

A;Molecule type: protein

A;Residues: 1-8 <HOL>

A;Cross-references: UNIPROT:P19988

C;Comment: Leucokinsins, a family of cephalomyotropic peptides, stimulate contractile act

C;Keywords: amidated carboxyl end; cephalomyotropic peptide; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.9%; Score 17; DB 2; Length 8;

Best Local Similarity 33.3%; Pred. No. 2.8e+05;

Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYHA 6  
| : : :  
Db 1 QSSFHS 6

RESULT 3  
NYPG7  
hypothalamic heptapeptide - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 09-Jul-2004  
C:Accession: A01417  
R;Chang, R.C.C.; Huang, W.Y.; Arimura, A.; Redding, T.W.; Coy, D.H.; Saffran, M.; Kong, H.; Chang, R.C.C.; Res. 13, 228-232, 1981  
A:Title: Isolation, structure and synthesis of a heptapeptide with in vitro ACTH-releasing activity  
A:Reference number: A01417; MUID:81213980; PMID:6263778  
A:Accession: A01417  
A:Molecule type: protein  
A:Residues: 1-7 <CHA>  
A:Cross-references: UNIPROT:P01153  
C:Superfamily: hypothalamic heptapeptide  
C:Keywords: hypothalamus

Query Match 29.1%; Score 16; DB 1; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHA 6  
| : :  
Db 3 YHS 5

RESULT 4  
PT0285  
Ig heavy chain CRD3 region (clone 4-100A) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0285  
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J<sub>H</sub> gene segments in the development of the B cell repertoire  
A:Reference number: PT0222; MUID:91108337; PMID:1899102  
A:Accession: PT0285  
A:Molecule type: DNA  
A:Residues: 1-9 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 29.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYH 5  
| : :  
Db 4 PTYH 7

RESULT 5  
S55696  
phosphoenolpyruvate carboxykinase - Trypanosoma brucei  
C:Species: Trypanosoma brucei  
C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S55696  
R;Hunt, M.; Koehler, P.  
Biochim. Biophys. Acta 1249, 15-22, 1995  
A:Title: Purification and characterization of phosphoenolpyruvate carboxykinase from Trypanosoma brucei  
A:Reference number: S55696; MUID:95284106; PMID:7766679  
A:Accession: S55696  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <HUN>  
A:Cross-references: UNIPROT:Q7M355

Query Match 29.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QPNYH 5  
| : :  
Db 1 QPIIH 5

RESULT 6  
A43405  
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC 3.1.3.1)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: A43405  
R;Ventura, F.; Rosa, J.L.; Ambrosio, S.; Palkis, S.J.; Bartrons, R.  
J. Biol. Chem. 267, 17939-17943, 1992  
A:Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence for a single protein with two distinct activities  
A:Reference number: A43405; MUID:92388154; PMID:1325453  
A:Accession: A43405  
A:Molecule type: protein  
A:Residues: 1-10 <VEN>  
A:Cross-references: UNIPROT:Q7M313  
C:Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phosphotransferase  
C:Keywords: phosphoric monoester hydrolase; phosphotransferase

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 3.8e+03;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QPNYH 5  
| : :  
Db 1 QVNH 5

RESULT 7  
S33844  
alpha-2-macroglobulin - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: S33844  
R;Warburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.  
Eur. J. Biochem. 214, 803-809, 1993  
A:Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary mammary gland  
A:Reference number: S33843; MUID:93307297; PMID:7686489  
A:Accession: S33844  
A:Molecule type: protein  
A:Residues: 1-10 <WAR>  
C:Superfamily: alpha-2-macroglobulin  
C:Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNY 4  
| : :  
Db 6 EPQY 9

RESULT 8  
PT0291  
Ig heavy chain CRD3 region (clone 4-115B) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0291  
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J<sub>H</sub> gene segments in the development of the B cell repertoire  
A:Reference number: PT0222; MUID:91108337; PMID:1899102  
A:Accession: PT0291  
A:Molecule type: DNA  
A:Residues: 1-10 <YAM>

A;Experimental source: B lymphocyte  
C;Keywords: heterotetramer; immunoglobulin

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYH 5  
| : |  
Db 4 PTYV 7

## RESULT 9

A59272  
peptide-N4- (N-acetyl-beta-glucosaminyl)asparagine amidase (EC 3.5.1.52) A, large chain  
N;Alternate names: peptide N-glycosidase  
C;Species: Prunus dulcis var. sativa (sweet almond)  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: A59272  
R;Altmann, F.; Paschinger, K.; Dalik, T.; Vorauer, K.  
Eur. J. Biochem. 252, 118-123, 1998  
A;Title: Characterisation of peptide-N4- (N-acetyl-beta-glucosaminyl)asparagine amidase A  
A;Reference number: A59272; MUID:98181894; PMID:9523720  
A;Accession: A59272  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <ALT>  
A;Cross-references: UNIPROT:P81899  
C;Keywords: hydrolase

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHA 6  
| : |  
Db 5 YHS 7

## RESULT 10

S23307  
neurokinin A - rainbow trout  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Aug-2004  
C;Accession: S23307  
R;Jensen, J.; Conlon, J.M.  
Eur. J. Biochem. 206, 659-664, 1992  
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod  
A;Reference number: S23186; MUID:92298992; PMID:1376687  
A;Accession: S23307  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <JEN>  
A;Cross-references: UNIPROT:P28500

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HAVN 8  
| : |  
Db 1 HKIN 4

## RESULT 11

S23186  
neurokinin A - Atlantic cod  
C;Species: Gadus morhua (Atlantic cod)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Aug-2004  
C;Accession: S23186  
R;Jensen, J.; Conlon, J.M.  
Eur. J. Biochem. 206, 659-664, 1992  
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod

A;Reference number: S23186; MUID:92298992; PMID:1376687  
A;Accession: S23186  
A;Molecule type: protein  
A;Residues: 1-10 <JEN>  
A;Cross-references: UNIPROT:P28500  
A;Experimental source: brain  
C;Function:  
A;Description: may play a physiological role in the regulation of cardiovascular and gas.  
A;Note: neurokinin A is derived by post-translational processing of preprotachykinin A  
C;Keywords: neuropeptide; amidated carboxyl end; tachykinin  
F;10/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HAVN 8  
| : |  
Db 1 HKIN 4

## RESULT 12

A38081  
amine oxidase (copper-containing) (EC 1.4.3.6) - yeast (*Pichia angusta*) (fragment)  
C;Species: *Pichia angusta*  
C;Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Apr-2000  
C;Accession: A38081  
R;Mu, D.; Jones, S.M.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Klinman, J.P.  
J. Biol. Chem. 267, 7979-7982, 1992  
A;Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine o.  
A;Reference number: A38081; MUID:92235001; PMID:1569055  
A;Accession: A38081  
A;Molecule type: protein  
A;Residues: 1-7 <MUA>  
C;Keywords: copper; oxidoreductase; quinoprotein; topaquinone  
F;4/Modified site: topaquinone (Tyr) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 7;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NYHAV 7  
| : |  
Db 3 NYEVV 7

## RESULT 13

A15398  
choline oxidase (EC 1.1.3.17) - *Alcaligenes* sp. (tentative sequence) (fragment)  
C;Species: *Alcaligenes* sp.  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: A15398  
R;Ohta-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.  
J. Biochem. 88, 197-203, 1980  
A;Title: Identification and properties of the prosthetic group of choline oxidase from *A*  
A;Reference number: A15398; MUID:81006769; PMID:6997283  
A;Accession: A15398  
A;Molecule type: protein  
A;Residues: 1-7 <OHT>  
A;Cross-references: UNIPROT:P16101  
C;Keywords: oxidoreductase

Query Match 27.3%; Score 15; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNYV 4  
| : |  
Db 3 PNH 5

## RESULT 14

A44960

neuropeptide Led-CC-I - Colorado potato beetle  
C;Species: Leptinotarsa decemlineata (Colorado potato beetle)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C;Accession: A44960  
R;Gaede, G.; Kellner, R.  
Peptides 10, 1287-1289, 1989  
A;Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
A;Reference number: A44960; MUID:90160053; PMID:2576128  
A;Accession: A44960  
A;Molecule type: protein  
A;Residues: 1-8 <GAE>  
A;Cross-references: UNIPROT:P04548  
C;Superfamily: adipokinetic hormone  
C;Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. NO. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNY 4  
||:  
Db 6 PNW 8

## RESULT 15

B44960  
neuropeptide Led-CC-II - Colorado potato beetle  
C;Species: Leptinotarsa decemlineata (Colorado potato beetle)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C;Accession: B44960  
R;Gaede, G.; Kellner, R.  
Peptides 10, 1287-1289, 1989  
A;Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
A;Reference number: A44960; MUID:90160053; PMID:2576128  
A;Accession: B44960  
A;Molecule type: protein  
A;Residues: 1-8 <GAE>  
A;Cross-references: UNIPROT:P04549  
C;Superfamily: adipokinetic hormone  
C;Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. NO. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNY 4  
||:  
Db 6 PNW 8

Search completed: May 19, 2005, 17:59:35  
Job time : 14 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:54:20 ; Search time 111 Seconds  
(without alignments)  
46.133 Million cell updates/sec

Title: US-09-867-159A-4  
Perfect score: 55  
Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 2548

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	23	41.8	10	2 Q8J0C2	Q8J0C2 encephalito
2	20	36.4	10	2 Q9UCQ8	Q9UCQ8 homo sapien
3	19	34.5	9	2 P90359	P90359 barley mild
4	19	34.5	10	2 Q6R7V4	Q6R7V4 carlia zuma
5	18	32.7	9	1 BS43 SERPL	P83375 serratia pl
6	18	32.7	9	2 Q08979	Q08979 mus musculus
7	18	32.7	10	2 Q7RGV8	Q7RGV8 plasmodium
8	18	32.7	10	2 P82438	P82438 nicotiana t
9	17	30.9	7	1 FAR4 PANRE	P41875 panagrellus
10	17	30.9	8	1 LCK6 LEUMA	P19988 leucophaea
11	17	30.9	10	2 Q9UCR0	Q9UCR0 homo sapien
12	17	30.9	10	2 Q7RSI4	Q7RSI4 plasmodium
13	17	30.9	10	2 Q8QE18	Q8QE18 human immun
14	16	29.1	7	1 HY7 PIG	P01153 sus scrofa
15	16	29.1	7	2 P72081	P72081 nocardia la
16	16	29.1	8	1 ALL1 CYDPO	P82152 cydia pomon
17	16	29.1	8	1 FAR1 PANRE	P41872 panagrellus
18	16	29.1	8	2 Q9UDZ4	Q9UDZ4 homo sapien
19	16	29.1	8	2 Q9UMH9	Q9UMH9 homo sapien
20	16	29.1	8	2 P83158	P83158 anabaena sp
21	16	29.1	9	1 FAR2 PANRE	P41873 panagrellus
22	16	29.1	9	2 Q7M355	Q7M355 trypanosoma
23	16	29.1	9	2 Q9TWX7	Q9TWX7 manduca sex
24	16	29.1	9	2 Q6XEV2	Q6XEV2 nectarinia
25	16	29.1	9	2 Q47410	Q47410 escherichia
26	16	29.1	9	2 Q9K4M6	Q9K4M6 staphylococ
27	16	29.1	10	1 PNAL PRUDU	P81899 prunus dulc
28	16	29.1	10	1 TKNB ONCMY	P28500 encorhynchu
29	16	29.1	10	2 Q9UCQ4	Q9UCQ4 homo sapien
30	16	29.1	10	2 Q7M313	Q7M313 bos taurus
31	16	29.1	10	2 Q9TS43	Q9TS43 sus scrofa

32	16	29.1	10	2 Q6XFW8	Q6XFW8 aethopyga b
33	16	29.1	10	2 Q86580	Q86580 simian para
34	15	27.3	7	1 CHOX ALCSP	P16101 alcaligenes
35	15	27.3	8	1 AKH GEOST	P84241 geotrupes s
36	15	27.3	8	1 AKH MELML	P84240 melolontha
37	15	27.3	8	1 AKH PACMA	P84242 pachnoda ma
38	15	27.3	8	1 HTFI PERAM	P04548 periplaneta
39	15	27.3	8	1 HTF2 TENNO	P04549 tenebrio mo
40	15	27.3	8	1 HTF ZOPRU	P67790 zophobas ru
41	15	27.3	8	1 HTF TENNO	Q15888 homo sapien
42	15	27.3	8	2 Q15888	Q15888 homo sapien
43	15	27.3	8	2 Q8IUB8	Q8IUB8 homo sapien
44	15	27.3	8	2 Q35792	Q35792 saccharomyc
45	15	27.3	8	2 Q6Y2F2	Q6Y2F2 citrus sine

ALIGNMENTS

RESULT 1  
Q8J0C2  
ID Q8J0C2 PRELIMINARY; PRT; 10 AA.  
AC Q8J0C2;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Hsp70 (Fragment).  
OS Encephalitozoon hellem.  
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
OX NCBI\_TaxID=27973;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PV6/95, and PV7/95;  
RX MEDLINE=22838663; PubMed=12958242;  
RA Haro M., Del Aguila C., Fenoy S., Henriques-Gil N.;  
RT "Intraspecies genotype variability of the microsporidian parasite  
J. Clin. Microbiol. 41:4166-4171(2003).  
RL EMBL; AV171238; AAN73415.1; -;  
DR EMBL; AV171239; AAN73417.1; -;  
FT NON TER 10 10  
SQ SEQUENCE 10 AA; 1100 MW; 107C0D472DD44DD4 CRC64;

Query Match 41.8%; Score 23; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 9.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYHAVN 8  
||| :||  
Db 2 PNANALN 8

RESULT 2  
Q9UCQ8  
ID Q9UCQ8 PRELIMINARY; PRT; 10 AA.  
AC Q9UCQ8;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE AUTOTAXIN (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92129337; PubMed=1733949;  
RA Stracke M.L., Krutzsch H.C., Unsworth E.J., Arestad A., Cioce V.,  
RA Schifmann E., Liotta L.A.;  
RT "Identification, purification, and partial sequence analysis of  
autotaxin, a novel motility-stimulating protein.";  
RL J. Biol. Chem. 267:2524-2529(1992).  
DR GO; GO:0006928; P:cell motility; NAS.

FT NON TER 1  
FT NON TER 10  
SQ SEQUENCE 10 AA; 1171 MW; 736F44577AF1B2CB CRC64;

Query Match 36.4%; Score 20; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNY 4  
Db 7 PNY 9

## RESULT 3

P90359 ID P90359 PRELIMINARY; PRT; 9 AA.  
AC P90359;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Polyprotein (Fragment).  
OS Barley mild mosaic virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Bymovirus.  
OX NCBI\_TaxID=12466;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Peerenboom E., Jacobi V., Cartwright E.J., Adams M.J., Steinbiss H.H.,  
RA Antoniw J.F.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X96970; CAA65658.1; -.  
KW Polyprotein.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 1177 MW; C40404473401F1B6 CRC64;

Query Match 34.5%; Score 19; DB 2; Length 9;  
Best Local Similarity 42.9%; Pred. No. 1.6e+06;  
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHAVN 8  
Db 1 POTHRLN 7

## RESULT 4

Q6R7V4 ID Q6R7V4 PRELIMINARY; PRT; 10 AA.  
AC Q6R7V4;  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).  
GN Name=GAPDH;  
OS Carlia zuma.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Scincoidea;  
OC Scincidae; Carlia.  
OX NCBI\_TaxID=260893;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dolman G., Phillips B.;  
RT "Single copy nuclear DNA markers characterized for comparative  
RT phylogeography in Australian wet tropics rainforest skinks.";  
RL Mol. Ecol. Notes 4:185-187(2004).  
DR EMBL; AY508912; AAS09890.1; -.  
FT NON TER 1  
FT NON TER 10  
SQ SEQUENCE 10 AA; 1171 MW; 9D0ABB2322C9C1EA CRC64;

Query Match 34.5%; Score 19; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.9e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYHAVNIV 10  
Db 2 NDHFVKLV 9

## RESULT 5

BS43\_SERPL ID BS43\_SERPL STANDARD; PRT; 9 AA.  
AC P83375;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Bacteriocin serracin P 43 kDa subunit (Fragment).  
OS Serratia plymuthica.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Serratia.  
OX NCBI\_TaxID=82996;  
RN [1]  
RP SEQUENCE, AND FUNCTION.  
RC STRAIN=J7;  
RX MEDLINE=22293561; PubMed=12406768;  
RA Jabrane A., Sabri A., Compare P., Jacques P., Vandenberghe I.,  
RA Van Beeumen J., Thonart P.;  
RT "Characterization of serracin P, a phage-tail-like bacteriocin, and  
RT its activity against *Erwinia amylovora*, the fire blight pathogen.";  
RL Appl. Environ. Microbiol. 68:5704-5710(2002).  
CC -!- FUNCTION: Major component of a prophage tail sheath (Probable).  
CC -!- FUNCTION: Antibacterial activity against Gram-negative bacterium  
CC *E. amylovora*.  
KW Antibiotic; Bacteriocin; Direct protein sequencing.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1095 MW; 1E66D412C871E1FB CRC64;

Query Match 32.7%; Score 18; DB 1; Length 9;  
Best Local Similarity 28.6%; Pred. No. 1.6e+06;  
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 YHAVNIV 10  
Db 3 HHGVRVL 9

## RESULT 6

O08979 ID O08979 PRELIMINARY; PRT; 9 AA.  
AC O08979;  
DT 01-JUL-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE AML1 protein (Fragment).  
GN Name=AML1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NMRI; TISSUE=Murine retrovirus induced tumor;  
RX MEDLINE=97332339; PubMed=9188573;  
RA Antoft H.W., Sorensen A.B., Bareil C., Schmidt J., Luz A.,  
RA Pedersen F.S.;  
RT "Stability of AML1 (core) site enhancer mutations in T-lymphomas  
RT induced by attenuated SL3-3 murine leukemia virus mutants.";  
RL J. Virol. 71:5080-5087(1997).  
DR EMBL; Y11802; CAA72496.1; -.  
FT NON TER 1  
FT NON TER 9  
SQ SEQUENCE 9 AA; 981 MW; 293E01E865A776D8 CRC64;

Query Match 32.7%; Score 18; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.6e+06;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 QPNVH 5
   ||: |
Db 3 QPSGH 7

RESULT 7
Q7RGV8 PRELIMINARY; PRT; 10 AA.
AC Q7RGV8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY04238;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Koolij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.B., Riedmuller S.B., Feidblyum I.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoalibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -I- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
KW EMBL; AABL01001270; EAA16068.1; -.
SQ SEQUENCE 10 AA; 1332 MW; F8601A30545B5051 CRC64;

Query Match 32.7%; Score 18; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 YHAVNI 9
   ||: |
Db 2 HHYNI 7

RESULT 8
P82438 PRELIMINARY; PRT; 10 AA.
AC P82438;
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 50 kDa cell wall protein (Fragment).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RT tobacco culture.";
RL Planta 0:0-0(2000).
CC -I- SUBCELLULAR LOCATION: CELL WALL.
CC -I- TISSUE SPECIFICITY: XYLEM.

DR GO; GO:0005618; C:cell wall; IEA.
KW Cell wall.
FT NON TER 10
SQ SEQUENCE 10 AA; 1126 MW; C68E32486AF77B46 CRC64;

Query Match 32.7%; Score 18; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 9.3e+03;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9
   ||: |
Db 2 PQYPXGNV 9

RESULT 9
FAR4_PANRE STANDARD; PRT; 7 AA.
AC P41875;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE FMRFamide-like neuropeptide PF4 (KPNFIRF-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=95232026; PubMed=7716079; DOI=10.1016/0196-9781(94)00162-Y;
RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P., Thim L.,
RA Kubiak T.M., Martin R.A., Geary F.G.;
RT "Isolation and preliminary biological characterization of
RT KPNFIRFamide, a novel FMRFamide-related peptide from the free-living
RT nematode, Panagrellus redivivus.";
RL Peptides 16:87-93(1995).
CC -I- FUNCTION: Myositive; induces a rapid concentration-dependent
CC muscle tension increase.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC family.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD RES 7 7 Phenylalanine amide.
SQ SEQUENCE 7 AA; 921 MW; 69D40059C4576350 CRC64;

Query Match 30.9%; Score 17; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4
   ||: |
Db 1 KPNF 4

RESULT 10
LCK6_LEUMA STANDARD; PRT; 8 AA.
AC P19988;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Leucokinin VI (u-VI).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RX MEDLINE=87052651; PubMed=2877794; DOI=10.1016/0742-8413(86)90077-0;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of Leucophaea maderae.";
```

RL Comp. Biochem. Physiol. 88C:27-30(1987).  
 CC -!- FUNCTION: This cephalomyotropic peptide stimulates contractile  
 CC activity of cockroach proctodeum (hindgut).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 DR FIR; JS0316; JS0316.  
 KW Amidation; Direct protein sequencing; Neuropeptide;  
 KW Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 Pyrrolidone carboxylic acid.  
 FT MOD\_RES 8 8 Glycine amide.  
 FT MOD\_RES 8 8 935 MW; 9D6365B1E9D5A5A6 CRC64;  
 SQ SEQUENCE 8 AA; 935 MW; 9D6365B1E9D5A5A6 CRC64;  
 Query Match 30.9%; Score 17; DB 1; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 1.6e+06;  
 Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QPNYHA 6  
 Db | : : : :  
 1 QSSFHS 6  
 RESULT 11  
 Q9UCR0 PRELIMINARY; PRT; 10 AA.  
 AC Q9UCR0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE AUTOTAXIN (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92129337; PubMed=1733949;  
 RA Stracke M.L., Krutzsch H.C., Unsworth E.J., Arestad A., Cioce V.,  
 RA Schiffrin E., Liotta L.A.;  
 RT "Identification, purification, and partial sequence analysis of  
 RT autotaxin, a novel motility-stimulating protein.";  
 RL J. Biol. Chem. 267:2524-2529(1992).  
 DR GO; GO:0005576; C:extracellular; IDA.  
 DR GO; GO:0030334; P:regulation of cell migration; IDA.  
 FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1193 MW; 3E01FB40544B19D7 CRC64;  
 Query Match 30.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+04;  
 Matches 2; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PNYHAVNI 9  
 Db | : : : :  
 2 PPENINL 9  
 RESULT 12  
 Q7RSI4 PRELIMINARY; PRT; 10 AA.  
 AC Q7RSI4;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=PY00373;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17XNI;  
 RX PubMed=42368865; DOI=10.1038/nature01099;  
 RA Carlton J.M., Anguolli S.V., Suh B.B., Kooij T.W., Peretea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,

RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,  
 RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,  
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,  
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 RT parasite Plasmodium yoelii yoelii.";  
 RL Nature 419:512-519(2002).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABL0100104; EAA15335.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 10 AA; 1279 MW; 11EBECB04B4B4B50 CRC64;  
 Query Match 30.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 14.3%; Pred. No. 1.5e+04;  
 Matches 1; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 YHAVNIV 10  
 Db | : : : :  
 4 YVVISVI 10  
 RESULT 13  
 Q8QE18 PRELIMINARY; PRT; 10 AA.  
 AC Q8QE18;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Truncated envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22961413; PubMed=14601597; DOI=10.1089/0892220322493139;  
 RA Masharsky A.E., Klimov N.A., Kozlov A.P.;  
 RT "Molecular cloning and analysis of full-length genome of HIV type 1  
 RT strains prevalent in countries of the former Soviet Union.";  
 RL AIDS Res. Hum. Retroviruses 19:933-939(2003).  
 DR EMBL; AF413977; AAL78427.1; -.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 KW Envelope protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1274 MW; A9FD6CCB544326D6 CRC64;  
 Query Match 30.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QPNY 4  
 Db | : : : :  
 4 QKNY 7  
 RESULT 14  
 HV7\_PIG STANDARD; PRT; 7 AA.  
 AC HV7\_PIG;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypothalamic heptapeptide.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;



```
[1]
RN SEQUENCE, AND SYNTHESIS.
RP MEDLINE=81213980; PubMed=6263778;
RX Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
RA Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT "Isolation, structure and synthesis of a heptapeptide with in vitro
RL ACTH-releasing activity from porcine hypothalamus.";
RM Horm. Metab. Res. 13:228-232 (1981).
DR PIR; A01417; NYPG7.
KW Direct protein sequencing.
SQ SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;

Query Match 29.1%; Score 16; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. NO. 1.6e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHA 6
Db 3 YHS 5

RESULT 15
P72081 PRELIMINARY; PRT; 7 AA.
AC P72081;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 3'-methylcephem hydroxylase (Fragment).
GN Name=cefF;
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009872; PubMed=7557411; DOI=10.1016/0378-1119(95)00308-S;
RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA Liras P.;
RT "Characterization of the cmh genes of Nocardia lactamdurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL; Z21682; CAA79797.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match 29.1%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. NO. 1.6e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HAV 7
Db 3 HAV 5

Search completed: May 19, 2005, 18:15:46
Job time : 115 secs
```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:47:19 ; Search time 141 Seconds  
(without alignments)  
27.430 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 465227

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	5	AAO20570 Cysteine
2	55	100.0	10	5	Abb98535 Cysteine
3	51	92.7	9	6	ABU11122 House dus
4	37	67.3	9	6	ABU11108 House dus
5	30	54.5	10	2	AA23224 Peptide d
6	28	50.9	8	4	AAE10569 Soybean p
7	28	50.9	9	8	ADK09720 Human eph
8	27	49.1	8	8	ADK09727 Human pap
9	27	49.1	9	8	ADK09741 Human pap
10	27	49.1	9	8	ADK09742 Human pap
11	27	49.1	9	8	ADK09740 Human pap
12	27	49.1	10	8	ADK09760 Human pap
13	26	47.3	8	5	ABP46729 Human Bly
14	26	47.3	8	7	ADG97556 scFV VHCD
15	26	47.3	9	6	ABR25375 Human can
16	26	47.3	9	6	ABR25590 Human can
17	26	47.3	9	6	ABR24431 Human can
18	26	47.3	9	6	ABR25198 Human can
19	26	47.3	10	2	AA96521 Hepatitis
20	26	47.3	10	6	ABR24938 Human can
21	26	47.3	10	6	ABR24674 Human can
22	26	47.3	10	6	ABR25279 Human can
23	26	47.3	10	6	ABR25474 Human can
24	26	47.3	10	6	ABR25523 Human can
25	26	47.3	10	6	ABR25760 Human can

26	26	47.3	10	6	ABR24521	Ab24521 Human can
27	26	47.3	10	6	ABR25108	Ab25108 Human can
28	26	47.3	10	6	ABR25710	Ab25710 Human can
29	25	45.5	8	4	AAU68238	Aau68238 Human Bre
30	25	45.5	8	6	ABB99503	Abb99503 Amino aci
31	25	45.5	8	6	ABG76122	Abg76122 Scrambled
32	25	45.5	8	8	ADR19346	Adr19346 TRP1 deri
33	25	45.5	9	2	AA773741	Aar773741 Antigen f
34	25	45.5	9	2	AA77563	Aar77563 HIV-B35-3
35	25	45.5	9	6	ABU11107	Abu11107 House dus
36	25	45.5	9	6	ABR05547	Ab05547 Human can
37	25	45.5	9	6	ABR05672	Ab05672 Human can
38	25	45.5	9	6	ABR05873	Ab05873 Human can
39	25	45.5	9	8	ADR11486	Adr11486 HLA-A2.1
40	25	45.5	10	1	AA82560	Ap82560 Guanine n
41	25	45.5	10	6	ABR05590	Ab05590 Human can
42	25	45.5	10	6	ABR05041	Ab05041 Human can
43	25	45.5	10	6	ABR05844	Ab05844 Human can
44	24	43.6	6	8	ADO37899	Ado37899 Binding p
45	24	43.6	6	8	ADO37682	Ado37682 Binding p

#### ALIGNMENTS

RESULT 1

AAO20570

ID AAO20570 standard; peptide; 10 AA.

XX AC AAO20570;

XX DT 02-JAN-2003 (first entry)

XX DE Cysteine protease epitope peptide region, SEQ ID No 4.

XX KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;

XX KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;

XX KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;

XX KW atopic eczema; epitope.

XX OS Dermatophagoides pteronyssinus.

XX PN WO200278736-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-FR001098.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PR 03-MAY-2001; 2001FR-00005929.

XX PR 29-MAY-2001; 2001US-00867159.

XX PA (ANTI-) ANTIALIS SARRL.

XX PI Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-750636/81.

XX DR Antiallergic compositions containing an anti-histamine, a histamine

XX PT synthesis inhibitor, and optionally an allergen or nucleic acid coding

XX PT for the allergen.

XX PS Claim 14; Page 11; 32pp; French.

XX CC The invention relates to antiallergic compositions containing an anti-

XX CC histamine, a histamine synthesis inhibitor, and optionally an allergen or

XX CC isolated nucleic acid molecule that has at least one polynucleotide

XX CC sequence coding for the allergen, together with a pharmaceutical carrier.

XX CC The pharmaceutical composition of the invention is useful as a non-

XX CC specific antiallergic treatment, and also useful in the treatment of

XX CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and

XX CC allergic and atopic eczema. This sequence represents a peptide of a

XX CC cysteine protease epitope region relating to the antiallergic

CC compositions of the invention  
XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00034;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10  
Db 1 QPNYHAVNIV 10

## RESULT 2

ABB98535  
ID ABB98535 standard; peptide; 10 AA.

XX  
AC ABB98535;

XX  
DT 13-DEC-2002 (first entry)

XX  
DE Cysteine protease epitope #2.

XX  
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunotherapy; allergen; allergic hypersensitivity reaction;  
KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
KW cysteine protease.

XX  
OS Dermatophagoides pteronyssinus.

XX  
PN FR2822709-A1.

XX  
PD 04-OCT-2002.

XX  
PF 03-MAY-2001; 2001FR-00005929.

XX  
PR 30-MAR-2001; 2001FR-00004370.

XX  
PA (ANTI-) ANTIALIS SARL.

XX  
PI Loria E, Terrasse G, Trehin Y;

XX  
DR WPI; 2002-735037/80.

XX  
PT Antiallergic composition, useful for preventing and treating e.g. asthma,  
PT rhinitis or eczema, containing at least two of allergen, antihistamine  
PT and histamine synthesis inhibitor.

XX  
PS Claim 8; Page 6; 33pp; French.

XX  
CC The present invention relates to an antiallergic pharmaceutical  
CC composition (I) comprising a pharmaceutical carrier containing an active  
CC agent combination of at least two of: an allergen; an antihistamine; and  
CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
CC rhinitis or allergic atopic eczema, in babies, children or adults. The  
CC present sequence is a peptide fragment (epitope) of cysteine protease  
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the  
CC invention

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00034;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10  
Db 1 QPNYHAVNIV 10

## RESULT 3

ABU11122

ID ABU11122 standard; peptide; 9 AA.

XX  
AC ABU11122;

DT 05-FEB-2003 (first entry)

XX  
DE House dust mite Der p1 antigen peptide #45.

XX  
KW House dust mite; Der p1 antigen; human CD8 cell epitope; allergy;  
KW immune response; atopic patient; CD8+ T-cell epitope; antiallergic.

XX  
OS Dermatophagoides pteronyssinus.

XX  
PN WO200281512-A1.

XX  
PD 17-OCT-2002.

XX  
PF 03-APR-2002; 2002WO-GB001534.

XX  
PR 06-APR-2001; 2001GB-00008752.

XX  
PA (ISIS-) ISIS INNOVATION LTD.

XX  
PI Ogg G, Seneviratne S;

XX  
DR WPI; 2003-058499/05.

XX  
PT New peptide fragments of the Der p1 antigen of the house dust mite  
PT Dermatophagoides pteronyssinus contain a human CD8+ T cell epitope and  
PT are useful to treat and prevent allergy to the major house dust mite  
PT allergen.

XX  
PS Disclosure; Page 32; 47pp; English.

XX  
CC The present invention relates to house dust mite (Dermatophagoides  
CC pteronyssinus) Der p1 antigen peptides containing human CD8 cell  
CC epitopes. The peptides of the invention are useful in the treatment of  
CC human or animal patients, particularly to raise an immune response to the  
CC Der p1 antigen. They are useful in the treatment and prevention of  
CC allergies to the major house dust mite antigen, and to monitor disease  
CC activity in atopic patients. ABU11078-ABU11146 represent house dust mite  
CC Der p1 antigen peptides containing CD8+ T-cell epitopes

XX  
SQ Sequence 9 AA;

Query Match 92.7%; Score 51; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNI 9  
Db 1 QPNYHAVNI 9

## RESULT 4

ABU11108  
ID ABU11108 standard; peptide; 9 AA.

XX  
AC ABU11108;

XX  
DT 05-FEB-2003 (first entry)

XX  
DE House dust mite Der p1 antigen peptide #31.

XX  
KW House dust mite; Der p1 antigen; human CD8 cell epitope; allergy;  
KW immune response; atopic patient; CD8+ T-cell epitope; antiallergic.

XX  
OS Dermatophagoides pteronyssinus.

XX  
PN WO200281512-A1.

XX  
PD 17-OCT-2002.

## RESULT 3

ABU11122

PF 03-APR-2002; 2002WO-GB001534.  
 XX  
 PR 06-APR-2001; 2001GB-00008752.  
 XX  
 PA (ISIS-) ISIS INNOVATION LTD.  
 XX  
 PI Ogg G, Seneviratne S;  
 XX  
 DR WPI; 2003-058499/05.  
 XX  
 XX New peptide fragments of the Der p1 antigen of the house dust mite  
 PT Dermatophagoides pteronyssinus contain a human CD8+ T cell epitope and  
 PT are useful to treat and prevent allergy to the major house dust mite  
 PT allergen.  
 XX  
 PS Disclosure; Page 31; 47pp; English.  
 XX  
 CC The present invention relates to house dust mite (Dermatophagoides  
 CC pteronyssinus) Der p1 antigen peptides containing human CD8 cell  
 CC epitopes. The peptides of the invention are useful in the treatment of  
 CC human or animal patients, particularly to raise an immune response to the  
 CC Der p1 antigen. They are useful in the treatment and prevention of  
 CC allergies to the major house dust mite antigen, and to monitor disease  
 CC activity in atopic patients. ABU11078-ABU11146 represent house dust mite  
 CC Der p1 antigen peptides containing CD8+ T-cell epitopes  
 XX  
 SQ Sequence 9 AA;  
 Query Match 67.3%; Score 37; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 YHAVNIV 10  
 DB 1 YHAVNIV 7  
 |||||  
 |||||  
 RESULT 5  
 AAY23224  
 ID AAY23224 standard; peptide; 10 AA.  
 AC  
 AC AAY23224;  
 XX  
 XX 25-AUG-1999 (first entry)  
 DT  
 DE Peptide derived from beta-1,2-xylosyltransferase.  
 XX  
 XX Beta-1,2-xylosyltransferase; beta 1,2-linked xylose; beta-linked mannose;  
 KW N-linked oligosaccharide; storage glycoprotein; allergenicity; soybean.  
 XX  
 OS Glycine max.  
 XX  
 PN WO9929835-A1.  
 XX  
 PD 17-JUN-1999.  
 XX  
 XX 08-DEC-1998; 98WO-US026047.  
 PF  
 PR 08-DEC-1997; 97US-0067932P.  
 PR 05-JAN-1998; 98US-0070418P.  
 XX  
 XX (UYAR-) UNIV ARKANSAS.  
 PA  
 XX Elbein AD, Bannon GA;  
 PI  
 XX WPI; 1999-385597/32.  
 DR  
 XX A Soybean-derived xylosyltransferase.  
 PT  
 PS Claim 3; Fig 4; 58pp; English.  
 XX  
 XX AAY23220-24 represent peptides released by beta-1,2-xylosyltransferase by  
 CC Endo lys C digestion. The specification describes a plant-derived beta

CC 1,2-xylosyltransferase, where the enzyme adds a beta 1,2-linked xylose to  
 CC the beta -linked mannose on the N-linked oligosaccharides of storage  
 CC glycoproteins. Xylose units on N-linked oligosaccharides may play a  
 CC critical role in allergenicity of plant derived glycoproteins and may  
 CC also be important in regulating the structure of the oligosaccharide  
 CC chains and the targeting of these proteins. Purification of a  
 CC xylosyltransferase is useful in order to study its properties and  
 CC specificities in the absence of interfering activities and possible  
 CC inhibitors. The antibody can be used to determine the localization of the  
 CC xylosyltransferase in suspension culture soybean cells and the  
 CC distribution and level of the enzyme in plants as well as its levels at  
 CC various stages of growth  
 XX  
 SQ Sequence 10 AA;  
 Query Match 54.5%; Score 30; DB 2; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 34;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 YHAVNI 9  
 DB 4 YHAINL 9  
 |||||  
 |||||  
 RESULT 6  
 AAE10569  
 ID AAE10569 standard; peptide; 8 AA.  
 XX  
 AC AAE10569;  
 XX  
 DT 10-DEC-2001 (first entry)  
 XX  
 XX Soybean peptide 3 to analyse A. thaliana beta 1,2-xylosyltransferase.  
 DE  
 DE Beta 1,2-xylosyltransferase; transgenic plant; medicament; glycoprotein;  
 KW soybean.  
 KW  
 XX Glycine max.  
 OS  
 XX WO200164901-A1.  
 PN  
 PD 07-SEP-2001.  
 XX  
 XX 02-MAR-2001; 2001WO-EP002352.  
 PF  
 XX 03-MAR-2000; 2000AT-00000355.  
 XX  
 XX (GLOE/) GLOESSL J.  
 PA  
 XX Gloessel J, Strasser R, Mucha J, Mach L, Altmann F, Wilson IB;  
 PI Steinkellner H;  
 PI  
 XX WPI; 2001-582160/65.  
 DR  
 XX Novel DNA molecule encoding beta 1,2-xylosyltransferase, useful for  
 PT producing transgenic plants and plant cells with increased efficiency in  
 PT producing glycoproteins.  
 PT  
 XX Example 1; Fig 1; 62pp; English.  
 PS  
 XX The invention relates to Arabidopsis thaliana beta 1,2-xylosyltransferase  
 CC plant protein and its cDNA molecule. Beta 1,2-xylosyltransferase protein  
 CC nucleic acid (PNA) molecule is useful for producing transgenic plants and  
 CC plant cells with increased efficiency in producing glycoproteins. The  
 CC invention also relates to a method for producing recombinant human  
 CC glycoproteins which is suitable for medical use. Beta 1,2-  
 CC xylosyltransferase DNA is useful for immobilisation on DNA microarrays,  
 CC e.g. for finding homologous sequences or for expression studies in plants  
 CC or non-vertebrate animals. The invention is also useful for inactivation,  
 CC suppression or over expression and production of beta 1,2-  
 CC xylosyltransferase. The present sequence is soybean peptide 3 which is  
 CC used for analysing Arabidopsis thaliana beta 1,2-xylosyltransferase  
 CC

SQ Sequence 8 AA;

Query Match 50.9%; Score 28; DB 4; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;  
 Matches 4; Conservative 1; Mismatches 0;

QY 4 YHAVN 8  
 |||:|  
 Db 4 YHAIN 8

RESULT 7  
 ADE85720  
 ID ADE85720 standard; peptide; 9 AA.  
 AC ADE85720;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Human EphA2 antibody Eph098B-233.152 VH CDR3 SEQ ID NO:24.  
 XX  
 KW cancer; hyperproliferative cell disease; EphA2 antibody;  
 KW EphA2 agonistic antibody; cytostatic; antiasthmatic; antipsooriatic;  
 KW antiinflammatory; vasotropic; respiratory; gene therapy;  
 KW metastatic cancer; asthma; psoriasis; inflammatory bowel disease;  
 KW smooth muscle restenosis; endothelial restenosis; Crohn's disease;  
 KW chronic obstructive pulmonary disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003094859-A2.  
 XX  
 PD 20-NOV-2003.  
 XX  
 PF 12-MAY-2003; 2003WO-US015044.  
 XX  
 PR 10-MAY-2002; 2002US-0379322P.  
 PR 14-OCT-2002; 2002US-0418213P.  
 PR 03-APR-2003; 2003US-0460507P.  
 XX  
 PA (MEDI-) MEDIMUNE INC.  
 XX  
 PI Kinch MS, Carles-Kinch K, Kiener P, Langermann S;  
 XX  
 DR WPI: 2004-012002/01.  
 DR N-PSDB; ADE85728.  
 XX  
 PT Treating cancer or a non-cancer hyperproliferative cell disease (e.g.  
 PT asthma, psoriasis, inflammatory bowel disease or restenosis) in a patient  
 PT comprises administering to the patient a therapeutic amount of an EphA2  
 PT antibody.  
 XX  
 PS Claim 54; SEQ ID NO 24; 173pp; English.

The present invention describes a method for treating cancer or a non-cancer hyperproliferative cell disease or disorder in a patient, which comprises administering to the patient a therapeutic amount of an EphA2 antibody (1) that is an EphA2 agonistic antibody, an EphA2 cancer cell phenotype inhibiting antibody, an exposed EphA2 epitope antibody, or an antibody that binds EphA2 with a K-off of less than 3 x 10<sup>-3</sup> s<sup>-1</sup>. Also described: (1) a pharmaceutical composition comprising a therapeutic amount of (1) and a pharmaceutical carrier; (2) a cell line that produces (1); (3) a hybridoma deposited with the ATCC accession number PTA-4572, PTA-4573 or PTA-4574; (4) an isolated nucleic acid comprising a nucleotide sequence encoding a light chain variable domain or a heavy chain variable domain of the EphA2 antibody; (5) a vector comprising the nucleic acid described above; (6) a host cell comprising the vector; (7) methods of identifying the EphA2 agonistic antibody or the EphA2 antibody that inhibits a cancer cell phenotype, that kills cancer cells having a cancer cell phenotype or that preferentially binds an EphA2 epitope exposed on cancer cells; and (8) a method of diagnosing, prognosing or monitoring the efficacy of therapy for cancer in a patient known to or suspected to have cancer. (1) has cytostatic, antiasthmatic,

CC antipsooriatic, antiinflammatory, vasotropic and respiratory activities, and can be used in gene therapy. The composition and methods are useful in managing, diagnosing, preventing or treating hyperproliferative cell diseases (i.e. metastatic cancer) or non-cancer hyperproliferative cell diseases or disorders, such as asthma, psoriasis, inflammatory bowel disease, smooth muscle restenosis, endothelial restenosis, Crohn's disease or chronic obstructive pulmonary disease. They may also be used for monitoring the efficacy of therapy for cancer in a patient known to or suspected to have cancer, and in screening for anti-cancer drugs. The present sequence is used in the exemplification of the present invention.

XX  
 SQ Sequence 9 AA;

Query Match 50.9%; Score 28; DB 8; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYHAVN 8  
 |||:|  
 Db 2 PRYHAMD 8

RESULT 8  
 ADK09727  
 ID ADK09727 standard; peptide; 8 AA.  
 XX  
 AC ADK09727;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human papillomavirus peptide #1782.  
 XX  
 KW pathogenic virus; alternative reading frame; antigenic determinant;  
 KW virucide; vaccine; therapeutic agent; infection; HPV.  
 XX  
 OS Human papillomavirus.  
 XX  
 PN WO2004011650-A2.  
 XX  
 PD 05-FEB-2004.  
 XX  
 PF 24-JUL-2003; 2003WO-EP008112.  
 XX  
 PR 24-JUL-2002; 2002AT-00001124.  
 PR 11-JUL-2003; 2003EP-00450171.  
 XX  
 PA (INTE-) INTERCELL AG.  
 XX  
 PI Mattner F, Schmidt W, Habel A;  
 XX  
 DR WPI: 2004-169243/16.  
 XX  
 PT New polypeptide encoded by an alternative reading frame of a pathogenic virus comprising an antigenic determinant, useful for treating or preventing an infection with the pathogenic virus.  
 XX  
 PS Claim 18; Page 193; 220pp; English.

This invention relates to a novel polypeptide encoded by an alternative reading frame of a pathogenic virus, where the polypeptide starts with a methionine amino acid residue, which comprises an antigenic determinant and more than 7 amino acid residues. The invention may be useful for the production of compounds with a virucide activity or the development of a vaccine. The polypeptide or its fragments may be useful as a therapeutic agent. It is also useful for the manufacture of a medicament for treating or preventing an infection with the pathogenic virus. The present sequence is that of a human papillomavirus (HPV) epitope peptide of the invention.

XX  
 SQ Sequence 8 AA;

Query Match 49.1%; Score 27; DB 8; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;

DE	Human papillomavirus peptide #1797.
XX	
KW	pathogenic virus; alternative reading frame; antigenic determinant;
KW	virucide; vaccine; therapeutic agent; infection; HPV.
XX	
OS	Human papillomavirus.
XX	
PN	WO2004011650-A2.
XX	
PD	05-FEB-2004.
XX	
PF	24-JUL-2003; 2003WO-EP008112.
XX	
PR	24-JUL-2002; 2002AT-00001124.
PR	11-JUL-2003; 2003EP-00450171.
XX	
PA	(INTE-) INTERCELL AG.
XX	
PI	Mattner F, Schmidt W, Habel A;
XX	
DR	WPI; 2004-169243/16.
XX	
PT	New polypeptide encoded by an alternative reading frame of a pathogenic
PT	virus comprising an antigenic determinant, useful for treating or
PT	preventing an infection with the pathogenic virus.
XX	
PS	Claim 18; Page 193; 220pp; English.
XX	
CC	This invention relates to a novel polypeptide encoded by an alternative
CC	reading frame of a pathogenic virus, where the polypeptide starts with a
CC	methionine amino acid residue, which comprises an antigenic determinant
CC	and more than 7 amino acid residues. The invention may be useful for the
CC	production of compounds with a virucide activity or the development of a
CC	vaccine. The polypeptide or its fragments may be useful as a therapeutic
CC	agent. It is also useful for the manufacture of a medicament for treating
CC	or preventing an infection with the pathogenic virus. The present
CC	sequence is that of a human papillomavirus (HPV) epitope peptide of the
CC	invention.
XX	
SQ	Sequence 9 AA;
Query Match	49.1%; Score 27; DB 8; Length 9;
Best Local Similarity	80.0%; Pred. No. 1.8e+06;
Matches	4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 QPNYH 5 
Dd	1 QPNYH 5 
RESULT 11	
ADK09740	
ID	ADK09740 standard; peptide; 9 AA.
XX	
AC	ADK09740;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human papillomavirus peptide #1795.
XX	
KW	pathogenic virus; alternative reading frame; antigenic determinant;
KW	virucide; vaccine; therapeutic agent; infection; HPV.
XX	
OS	Human papillomavirus.
XX	
PN	WO2004011650-A2.
XX	
PD	05-FEB-2004.
XX	
PF	24-JUL-2003; 2003WO-EP008112.
XX	
PR	24-JUL-2002; 2002AT-00001124.
PR	11-JUL-2003; 2003EP-00450171.

```

XX (INTE-) INTERCELL AG.
XX Mattner F, Schmidt W, Habel A;
XX WPI; 2004-169243/16.
XX New polypeptide encoded by an alternative reading frame of a pathogenic
XX virus comprising an antigenic determinant, useful for treating or
XX preventing an infection with the pathogenic virus.
XX Claim 18; Page 193; 220pp; English.
XX This invention relates to a novel polypeptide encoded by an alternative
XX reading frame of a pathogenic virus, where the polypeptide starts with a
XX methionine amino acid residue, which comprises an antigenic determinant
XX and more than 7 amino acid residues. The invention may be useful for the
XX production of compounds with a virucide activity or the development of a
XX vaccine. The polypeptide or its fragments may be useful as a therapeutic
XX agent. It is also useful for the manufacture of a medicament for treating
XX or preventing an infection with the pathogenic virus. The present
XX sequence is that of a human papillomavirus (HPV) epitope peptide of the
XX invention.
XX Query Match 49.1%; Score 27; DB 8; Length 9;
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 QPNYH 5
XX 4 QPRYH 8
XX
XX RESULT 12
XX ADK09760
XX ID ADK09760 standard; peptide; 10 AA.
XX AC ADK09760;
XX
XX 06-MAY-2004 (first entry)
XX Human papillomavirus peptide #1815.
XX pathogenic virus; alternative reading frame; antigenic determinant;
XX virucide; vaccine; therapeutic agent; infection; HPV.
XX Human papillomavirus.
XX WO2004011650-A2.
XX
XX 05-FEB-2004.
XX
XX 24-JUL-2003; 2003WO-EP008112.
XX
XX 24-JUL-2002; 2002AT-00001124.
XX 11-JUL-2003; 2003EP-00450171.
XX (INTE-) INTERCELL AG.
XX Mattner F, Schmidt W, Habel A;
XX WPI; 2004-169243/16.
XX New polypeptide encoded by an alternative reading frame of a pathogenic
XX virus comprising an antigenic determinant, useful for treating or
XX preventing an infection with the pathogenic virus.
XX Claim 18; Page 193; 220pp; English.
XX This invention relates to a novel polypeptide encoded by an alternative
XX reading frame of a pathogenic virus, where the polypeptide starts with a
XX methionine amino acid residue, which comprises an antigenic determinant
XX and more than 7 amino acid residues. The invention may be useful for the
XX production of compounds with a virucide activity or the development of a
XX vaccine. The polypeptide or its fragments may be useful as a therapeutic
XX agent. It is also useful for the manufacture of a medicament for treating
XX or preventing an infection with the pathogenic virus. The present
XX sequence is that of a human papillomavirus (HPV) epitope peptide of the
XX invention.
XX Query Match 49.1%; Score 27; DB 8; Length 9;
XX Best Local Similarity 80.0%; Pred. No. 1.8e+06;
XX Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 QPNYH 5
XX 4 QPRYH 8
XX
XX RESULT 13
XX ABP46729
XX ID ABP46729 standard; peptide; 8 AA.
XX AC ABP46729;
XX
XX 19-AUG-2002 (first entry)
XX Human BlyS binding scFv VH CDR3 SEQ ID 2740.
XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
XX tumour necrosis factor; B cell proliferation; B cell differentiation;
XX immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
XX antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
XX common variable immunodeficiency; acquired immunodeficiency syndrome.
XX Homo sapiens.
XX WO200202641-A1.
XX
XX 10-JAN-2002.
XX
XX 15-JUN-2001; 2001WO-US019110.
XX
XX 16-JUN-2000; 2000US-0212210P.
XX 17-OCT-2000; 2000US-0240816P.
XX 16-MAR-2001; 2001US-0276248P.
XX 21-MAR-2001; 2001US-0277379P.
XX 25-MAY-2001; 2001US-0293499P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX WPI; 2002-114799/15.
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
XX diagnosis and treatment of cancers and immune disorders.
XX Claim 2; Page 3042; 3148pp; English.
XX This invention describes novel antibodies that immunospecifically bind to
XX B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
XX tumour necrosis factor (TNF) super family and induces B cell
XX proliferation and differentiation. The antibodies of the invention have
XX cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
XX antirheumatic and antiAIDS activity and can be used in vaccines to
XX inhibit the expression and activity of BlyS. The antibodies bind to BlyS
XX and so may be used to detect and quantitate the presence of BlyS in
XX biological samples and may be used in this way to diagnose disease
XX associated with aberrant expression of BlyS. They may also be

```



CC administered to treat diseases associated with aberrant BlyS expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention

XX Sequence 8 AA;

Query Match 47.3%; Score 26; DB 5; Length 8;  
 Best Local Similarity 37.5%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYHAVNI 9  
 Db 1 PSYHYMDV 8

RESULT 14

ID ADG97556

ADG97556 standard; peptide; 8 AA.

XX AC ADG97556;

DT 11-MAR-2004 (first entry)

DE scFV VHCDR3 peptide that immunospecifically binds BlyS SeqID 2740.

XX antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;  
 KW B cell proliferation; differentiation; scFv; myasthenia gravis;  
 KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
 KW carcinoma; lymphoma; antiarthritis; antiarthritis; neuroprotective;  
 KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

XX WO2003055979-A2.

XX 10-JUL-2003.

XX 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

XX 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
 PT (BlyS), useful for detecting and treating diseases or disorders e.g.  
 PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 2740; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
 CC to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to  
 CC chromosome 13q34 and encodes a protein that is a member of the tumour  
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
 CC proliferation and differentiation. Specifically, it refers to single  
 CC chain antibody molecules (scFvs) derived, preferably, from the variable  
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
 CC fragment thereof, of either human, murine, rat or monkey BlyS. The  
 CC present invention refers to the use of such antibodies in various methods  
 CC for the detection, diagnosis and prognosis of diseases related to the  
 CC aberrant expression or inappropriate function of BlyS or its receptor. As  
 CC such, these compositions are useful for identifying immune disorders  
 CC including myasthenia gravis and multiple sclerosis, inflammatory  
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
 CC as AIDS and proliferative disorders including leukaemia, carcinoma and

CC lymphoma. Accordingly, they can be described as exhibiting various  
 CC activities such as antirheumatic, antiarthritic, neuroprotective,  
 CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
 CC peptide sequence is a single chain antibody variable heavy CDR3 peptide  
 CC that immunospecifically binds BlyS of the invention.

XX Sequence 8 AA;

Query Match 47.3%; Score 26; DB 7; Length 8;  
 Best Local Similarity 37.5%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYHAVNI 9  
 Db 1 PSYHYMDV 8

RESULT 15

ABR25375  
 ID ABR25375 standard; peptide; 9 AA.

XX AC ABR25375;

XX 19-MAY-2003 (first entry)

XX Human cancer-related protein 185P3C3 HLA peptide #1010.

XX Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.

XX Homo sapiens.

XX WO200283921-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US011654.

XX 10-APR-2001; 2001US-0282739P.

XX 10-APR-2001; 2001US-0283112P.

XX 25-APR-2001; 2001US-0286630P.

XX (AGEN-) AGENSYS INC.

XX Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;  
 PI Morrison K, Morrison RK, Raitano AB;

XX WPI; 2003-075555/07.

XX New composition comprising a substance that modulates the structure of  
 PT proteins and polynucleotides, useful for therapeutic, prognostic and  
 PT diagnostic reagents for eliciting cellular or humoral immune response in  
 PT cancer patients.

XX Claim 13; Page 380; 1021pp; English.

XX The present invention relates to novel human cancer-related genes and  
 CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
 CC proteins are useful for eliciting a humoral or cellular immune response.  
 CC The genes are useful as probes and primers for the amplification and/or  
 CC detection of genes, mRNAs or their fragments, as reagents for the  
 CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
 CC directing the expression of the protein, as tools for modulating or  
 CC inhibiting the expression of genes and/or translation of transcripts, and  
 CC as therapeutic agents. The proteins and peptides are useful as  
 CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
 CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention

XX Sequence 9 AA;

Query Match 47.3%; Score 26; DB 6; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 1.8e+06;

Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 BNYHAVNI 9

Db 2 PDFHSENL 9

Search completed: May 19, 2005, 17:59:10  
Job time : 146 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:59:41 ; Search time 89.5 Seconds  
(without alignments)  
37.375 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 185062

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	10	US-09-867-159A-4
2	30	54.5	10	9	US-09-748-578-5
3	30	54.5	10	14	US-10-411-905-5
4	28	50.9	8	16	US-10-220-467A-2
5	28	50.9	9	15	US-10-436-782-24
6	26	47.3	8	10	US-09-880-748-2740
7	26	47.3	8	15	US-10-293-418-2740
8	26	47.3	10	9	US-09-851-138-204
9	25	45.5	8	10	US-09-988-493-162
10	25	45.5	8	13	US-10-007-363-4
11	25	45.5	8	16	US-10-807-553-4
12	25	45.5	9	17	US-10-820-467-236
13	24	43.6	6	17	US-10-699-114-742

14	24	43.6	6	17	US-10-699-114-959	Sequence 959, App
15	24	43.6	6	17	US-10-806-924-111	Sequence 11, Appl
16	24	43.6	6	17	US-10-806-924-705	Sequence 705, App
17	24	43.6	7	14	US-10-052-578-164	Sequence 164, App
18	24	43.6	7	14	US-10-052-578-210	Sequence 210, App
19	24	43.6	7	14	US-10-053-520-164	Sequence 164, App
20	24	43.6	7	14	US-10-053-520-210	Sequence 210, App
21	24	43.6	7	14	US-10-053-4988-164	Sequence 164, App
22	24	43.6	7	14	US-10-053-4988-210	Sequence 210, App
23	24	43.6	7	15	US-10-258-146A-22	Sequence 22, Appl
24	24	43.6	7	15	US-10-258-146A-68	Sequence 68, Appl
25	24	43.6	7	15	US-10-328-953-167	Sequence 167, App
26	24	43.6	7	15	US-10-328-953-213	Sequence 213, App
27	24	43.6	7	16	US-10-258-144-57	Sequence 57, Appl
28	24	43.6	7	16	US-10-258-144-103	Sequence 103, App
29	24	43.6	9	9	US-09-826-177-56	Sequence 56, Appl
30	24	43.6	9	15	US-10-428-335-142	Sequence 142, App
31	23	41.8	6	17	US-10-699-114-760	Sequence 760, App
32	23	41.8	6	17	US-10-699-114-1053	Sequence 1053, App
33	23	41.8	6	17	US-10-806-924-105	Sequence 105, App
34	23	41.8	6	17	US-10-806-924-723	Sequence 723, App
35	23	41.8	7	15	US-10-285-394-289	Sequence 289, App
36	23	41.8	7	15	US-10-601-837-176	Sequence 176, App
37	23	41.8	9	9	US-09-826-177-40	Sequence 40, Appl
38	23	41.8	9	9	US-09-826-177-72	Sequence 72, Appl
39	23	41.8	9	9	US-09-826-177-74	Sequence 74, Appl
40	23	41.8	9	9	US-09-826-177-76	Sequence 76, Appl
41	23	41.8	9	9	US-09-826-177-84	Sequence 84, Appl
42	23	41.8	9	14	US-10-254-446A-109	Sequence 109, App
43	23	41.8	9	16	US-10-415-014-106	Sequence 106, App
44	23	41.8	9	16	US-10-415-014-219	Sequence 219, App
45	23	41.8	9	16	US-10-415-014-414	Sequence 414, App

#### ALIGNMENTS

#### RESULT 1

US-09-867-159A-4  
; Sequence 4, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antialis and at least one anti-histamine compound  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Comprises epitope from cystine protease.

US-09-867-159A-4

Query Match 100.0%; Score 55; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10

|||||

Db 1 QPNYHAVNIV 10

```
RESULT 2
US-09-748-578-5
; Sequence 5, Application US/09748578
; Patent No. US20010016344A1
; GENERAL INFORMATION:
; APPLICANT: Elbein, Alan D.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof
; FILE REFERENCE: D6063/D
; CURRENT APPLICATION NUMBER: US/09/748,578
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/207,223
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: soybean
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of a peptide released by Endo
; OTHER INFORMATION: lys C digestion of purified xylosyltransferase.
US-09-748-578-5
Query Match 54.5%; Score 30; DB 9; Length 10;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVNI 9
DB 4 YHAINL 9

RESULT 3
US-10-411-905-5
; Sequence 5, Application US/10411905
; Publication No. US20030166012A1
; GENERAL INFORMATION:
; APPLICANT: Elbein, Alan D.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof
; FILE REFERENCE: D6063/D2
; CURRENT APPLICATION NUMBER: US/10/411,905
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: US 09/748,578
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: soybean
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of a peptide released
; OTHER INFORMATION: by Endo lys C digestion of purified
; OTHER INFORMATION: xylosyltransferase.
US-10-411-905-5
Query Match 54.5%; Score 30; DB 14; Length 10;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVNI 9
DB 4 YHAINL 9

RESULT 4
US-10-220-467A-2
; Sequence 2, Application US/10220467A
; Publication No. US20040121325A1
; GENERAL INFORMATION:
; APPLICANT: Glossl Prof., Josef
; TITLE OF INVENTION: Beta 1, 2-Xylosyltransferase-Gene from Arabidopsis
; FILE REFERENCE: SONN:019US
```

```
; CURRENT APPLICATION NUMBER: US/10/220,467A
; CURRENT FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: A 355/2000
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 8
; TYPE: PRT
; ORGANISM: soybean
US-10-220-467A-2
Query Match 50.9%; Score 28; DB 16; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.3e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVN 8
DB 4 YHAIN 8

RESULT 5
US-10-436-782-24
; Sequence 24, Application US/10436782
; Publication No. US20040028685A1
; GENERAL INFORMATION:
; APPLICANT: Kinch, Michael
; APPLICANT: Carles-Kinch, Kelly
; APPLICANT: Kiener, Peter
; APPLICANT: Langermann, Solomon
; TITLE OF INVENTION: EphA2 Monoclonal Antibodies and Methods of Use Thereof
; FILE REFERENCE: 10271-097
; CURRENT APPLICATION NUMBER: US/10/436,782
; CURRENT FILING DATE: 2003-05-12
; PRIOR APPLICATION NUMBER: 60/379,322
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/418,213
; PRIOR FILING DATE: 2002-10-14
; PRIOR APPLICATION NUMBER: 60/418,213
; PRIOR FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-436-782-24
Query Match 50.9%; Score 28; DB 15; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.3e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYHAVN 8
DB 2 PRYHAMD 8

RESULT 6
US-09-880-748-2740
; Sequence 2740, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
```

TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES  
AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC  
AGENTS

```

RESULT 9
US-09-988-493-162
; Sequence 162, Application US/09988493
; Publication No. US20030064419A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: O'Hare, Michael John
; APPLICANT: Page, Martin John
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Waterfield, Michael Derek
; TITLE OF INVENTION: Proteins, Genes, and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Breast Cancer
; FILE REFERENCE: 2543-1-024
; CURRENT APPLICATION NUMBER: US/09/988,493
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/GB01/01219
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: GB 0006695.1
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: GB 0007265.2
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 8
; TYPE: PRT
; ORGANISM: homo sapien
US-09-988-493-162

```

Query Match 45.5%; Score 25; DB 10; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NYHAV 7  
|:|:|  
Db 2 NFHAV 6

RESULT 10  
US-10-007-363-4  
; Sequence 4, Application US/10007363  
; Publication No. US20020168354A1  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: pseudo-epsilon RACK Peptide Composition  
; TITLE OF INVENTION: and Method for Protection Against Tissue Damage Due to Ischemia  
; FILE OF INVENTION: Ischemia  
; FILE REFERENCE: 58600-8209 US00  
; CURRENT APPLICATION NUMBER: US/10/007,363  
; CURRENT FILING DATE: 2002-11-09  
; PRIOR APPLICATION NUMBER: US 60/247,830  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: scrambled pseudo-epsilon RACK octapeptide  
US-10-007-363-4

Query Match 45.5%; Score 25; DB 13; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9  
|:|:|  
Db 1 PDYHDAGI 8

RESULT 11  
US-10-807-553-4  
; Sequence 4, Application US/10807553  
; Publication No. US20040186055A1  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: pseudo-epsilon RACK Peptide Composition  
; TITLE OF INVENTION: and Method for Protection Against Tissue Damage Due to Ischemia  
; FILE REFERENCE: 58600-8209 US00  
; CURRENT APPLICATION NUMBER: US/10/807,553  
; CURRENT FILING DATE: 2004-03-22  
; PRIOR APPLICATION NUMBER: US/10/007,363  
; PRIOR FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: US 60/247,830  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: scrambled pseudo-epsilon RACK octapeptide  
US-10-807-553-4

Query Match 45.5%; Score 25; DB 16; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9  
|:|:|  
Db 1 PDYHDAGI 8

RESULT 12  
US-10-820-467-236  
; Sequence 236, Application US/10820467  
; Publication No. US20050054053A1  
; GENERAL INFORMATION:  
; APPLICANT: Aguinaldo, Anna Marie  
; APPLICANT: Beyna, Amelia Joy  
; APPLICANT: Cho, Ho Sung  
; APPLICANT: Desjarlais, John Rudolph  
; APPLICANT: Marshall, Shannon Alicia  
; APPLICANT: Muchhal, Umesh  
; APPLICANT: Villegas, Michael Francis Aquino  
; APPLICANT: Zhukovsky, Eugene  
; APPLICANT: Quesenberry, Michael Stephen  
; TITLE OF INVENTION: INTERFERON VARIANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: A-71431-4  
; CURRENT APPLICATION NUMBER: US/10/820,467  
; CURRENT FILING DATE: 2004-03-30  
; PRIOR APPLICATION NUMBER: US 60/477,246  
; PRIOR FILING DATE: 2003-06-10  
; PRIOR APPLICATION NUMBER: US 60/415,541  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: US 60/489,725  
; PRIOR FILING DATE: 2003-07-24  
; PRIOR APPLICATION NUMBER: US 10/676,705  
; PRIOR FILING DATE: 2003-09-30  
; NUMBER OF SEQ ID NOS: 274  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 236  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-10-820-467-236

Query Match 45.5%; Score 25; DB 17; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYHAVN 8  
|:|:|  
Db 3 NFHYVN 8

RESULT 13  
US-10-699-114-742  
; Sequence 742, Application US/10699114  
; Publication No. US20050042623A1  
; GENERAL INFORMATION:  
; APPLICANT: Dana Ault-Riche  
; APPLICANT: Bruce Atkinson  
; APPLICANT: Krishnanand Kumble  
; APPLICANT: Lynne Jersaltis  
; APPLICANT: Gizette Sperinde  
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH  
; TITLE OF INVENTION: USING THE SYSTEMS  
; FILE REFERENCE: 25885-1759  
; CURRENT APPLICATION NUMBER: US/10/699,114  
; CURRENT FILING DATE: 2003-10-30  
; PRIOR APPLICATION NUMBER: 60/423,018  
; PRIOR FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/422,923  
; PRIOR FILING DATE: 2002-10-30  
; NUMBER OF SEQ ID NOS: 1094  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 742  
; LENGTH: 6

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-742

Query Match      43.6%; Score 24; DB 17; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.3e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QPNYH 5
Db      2 EPGYH 6

RESULT 14
US-10-699-114-959
; Sequence 959, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METHODS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 959
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-959

Query Match      43.6%; Score 24; DB 17; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.3e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QPNYH 5
Db      1 EPGYH 5

RESULT 15
US-10-806-924-11
; Sequence 11, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
```

US-10-806-924-11

Query Match 43.6%; Score 24; DB 17; Length 6;  
Best Local Similarity 60.0%; Pred. No. 1.3e+06;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
Db 1 EPGYH 5

Search completed: May 19, 2005, 18:19:59  
Job time : 90.5 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 17:53:34 ; Search time 29.5 Seconds  
(without alignments)  
25.305 Million cell updates/sec

Title: US-09-867-159A-4  
Perfect score: 55  
Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 115750

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents RA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	54.5	10	3	US-09-207-223-5
2	30	54.5	10	4	US-09-748-578-5
3	26	47.3	10	3	US-08-836-075A-204
4	25	45.5	9	1	US-08-615-181-108
5	25	45.5	10	1	US-08-212-433A-33
6	25	45.5	10	3	US-08-716-256-33
7	25	45.5	10	5	PCT-US95-03239-33
8	25	45.5	10	6	5436320-3
9	25	45.5	10	6	5436320-3
10	23	41.8	6	4	US-08-877-605-181
11	23	41.8	6	4	US-08-877-605-220
12	22	40.0	4	4	US-09-057-363C-18
13	22	40.0	4	1	US-09-265-107-18
14	22	40.0	7	1	US-08-092-110A-3
15	22	40.0	7	1	US-08-273-474-3
16	22	40.0	7	3	US-08-405-647B-14
17	22	40.0	7	3	US-08-935-100-3
18	22	40.0	7	3	US-09-147-933-22
19	22	40.0	7	3	US-08-985-499-14
20	22	40.0	7	5	PCT-US94-07881-3
21	22	40.0	7	5	PCT-US96-03180-14
22	22	40.0	8	3	US-08-947-965-62
23	22	40.0	8	3	US-08-444-818-404
24	22	40.0	9	2	US-08-318-856A-34
25	22	40.0	10	3	US-08-836-075A-195
26	21	38.2	4	1	US-08-456-424-79
27	21	38.2	4	1	US-08-456-424-80

28	21	38.2	5	2	US-08-558-823-12	Sequence 12, Appl
29	21	38.2	5	3	US-08-604-991-14	Sequence 14, Appl
30	21	38.2	5	3	US-09-363-639-14	Sequence 14, Appl
31	21	38.2	6	2	US-08-428-131-3	Sequence 3, Appli
32	21	38.2	6	2	US-08-558-823-15	Sequence 15, Appl
33	21	38.2	6	2	US-08-558-823-16	Sequence 16, Appl
34	21	38.2	6	2	US-08-310-912A-104	Sequence 104, App
35	21	38.2	6	3	US-08-893-534A-44	Sequence 44, Appl
36	21	38.2	6	3	US-08-841-089-104	Sequence 104, App
37	21	38.2	6	3	US-09-078-596-3	Sequence 3, Appli
38	21	38.2	6	3	US-08-939-853A-28	Sequence 28, Appl
39	21	38.2	6	3	US-08-996-679-44	Sequence 44, Appl
40	21	38.2	6	3	US-09-115-395-19	Sequence 19, Appl
41	21	38.2	6	3	US-09-301-085-104	Sequence 104, App
42	21	38.2	6	3	US-09-507-102-44	Sequence 44, Appl
43	21	38.2	6	3	US-09-250-059-34	Sequence 34, Appl
44	21	38.2	6	3	US-09-248-074-34	Sequence 34, Appl
45	21	38.2	6	3	US-09-187-859-60	Sequence 60, Appl

## ALIGNMENTS

## RESULT 1

US-09-207-223-5

; Sequence 5, Application US/09207223

; Patent No. 6168937

; GENERAL INFORMATION:

; APPLICANT: Elbein, Alan D.

; APPLICANT: Bannon, Gary A.

; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof

; FILE REFERENCE: D6063

; CURRENT APPLICATION NUMBER: US/09/207,223

; CURRENT FILING DATE: 1998-12-08

; EARLIER APPLICATION NUMBER: US 60/067,932

; EARLIER FILING DATE: 1997-12-08

; NUMBER OF SEQ ID NOS: 7

; SEQ ID NO 5

; LENGTH: 10

; TYPE: PRT

; ORGANISM: soybean

; FEATURE:

; OTHER INFORMATION: Amino acid sequence of a peptide released by Endo lys C

; OTHER INFORMATION: digestion of purified xylosyltransferase.

US-09-207-223-5

Query Match 54.5%; Score 30; DB 3; Length 10;

Best Local Similarity 66.7%; Pred. No. 15;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YHAVNI 9

Db 4 YHAINL 9

## RESULT 2

US-09-748-578-5

; Sequence 5, Application US/09748578

; Patent No. 6593462

; GENERAL INFORMATION:

; APPLICANT: Elbein, Alan D.

; APPLICANT: Bannon, Gary A.

; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof

; FILE REFERENCE: D6063/D

; CURRENT APPLICATION NUMBER: US/09/748,578

; CURRENT FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/207,223

; PRIOR FILING DATE: 1998-12-08

; NUMBER OF SEQ ID NOS: 7

; SEQ ID NO 5

; LENGTH: 10

; TYPE: PRT

; ORGANISM: soybean

```
/ FEATURE:
/ OTHER INFORMATION: Amino acid sequence of a peptide released by Endo
/ OTHER INFORMATION: lys C digestion of purified xylosyltransferase.
US-09-748-578-5

Query Match      54.5%; Score 30; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVNI 9
   |||:|
Db 4 YHAINL 9

RESULT 3
US-08-836-075A-204
/ Sequence 204, Application US/08836075A
/ Patent No. 6180768
/ GENERAL INFORMATION:
/ APPLICANT: MAERTENS, GEBRT
/ APPLICANT: STUYVER, LIEVEN
/ TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
/ TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
/ TITLE OF INVENTION: AGENTS
/ NUMBER OF SEQUENCES: 207
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: ARNOLD, WHITE & DURKEE
/ STREET: P.O. BOX 4433
/ CITY: HOUSTON
/ STATE: TEXAS
/ COUNTRY: USA
/ ZIP: 77210-4433
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Microsoft Word 6.0 / ASCII text output
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/836,075A
/ FILING DATE: 21 Apr 1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/EP95/04155
/ FILING DATE: 23 Oct 1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 94870166.9
/ FILING DATE: 21 Oct 1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95870076.7
/ FILING DATE: 28 Jun 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KAMMERER, PATRICIA A.
/ REGISTRATION NUMBER: 29,775
/ REFERENCE/DOCKET NUMBER: INNS:004
/ INFORMATION FOR SEQ ID NO: 204:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-836-075A-204

Query Match      47.3%; Score 26; DB 3; Length 10;
Best Local Similarity 57.1%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QPNYHAV 7
   :|||
Db 1 RPKYQV 7

RESULT 4
US-08-615-181-108
/ Sequence 108, Application US/08615181
```

```
/ Patent No. 5756666
/ GENERAL INFORMATION:
/ APPLICANT: MASAFUMI, TAKIGUCHI
/ APPLICANT: MIWA, KIYOSHI
/ TITLE OF INVENTION: PEPTIDES CAPABLE OF INDUCING IMMUNE
/ TITLE OF INVENTION: RESPONSE TO HIV AND ANTI-AIDS AGENT FOR PREVENTING AND
/ TITLE OF INVENTION: CURING AIDS
/ NUMBER OF SEQUENCES: 115
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
/ ADDRESSEE: P.C.
/ STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
/ CITY: ARLINGTON
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/615,181
/ FILING DATE: 04-APR-1996
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP94/01756
/ FILING DATE: 19-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 261302/1993
/ FILING DATE: 19-OCT-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: OBLON, NORMAN F.
/ REGISTRATION NUMBER: 24,618
/ REFERENCE/DOCKET NUMBER: 10-796-0 PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-413-3000
/ TELEFAX: 703-413-2220
/ INFORMATION FOR SEQ ID NO: 108:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
/ US-08-615-181-108

Query Match      45.5%; Score 25; DB 1; Length 9;
Best Local Similarity 62.5%; Pred. No. 4,1e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9
   |:|||
Db 2 PSGHAVGI 9

RESULT 5
US-08-212-433A-33
/ Sequence 33, Application US/08212433A
/ Patent No. 5538897
/ GENERAL INFORMATION:
/ APPLICANT: Yates, III, John R.
/ APPLICANT: Eng, James K.
/ TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION
/ TITLE OF INVENTION: PATTERNS OF PEPTIDES TO IDENTIFY AMINO ACID SEQUENCES IN
/ TITLE OF INVENTION: DATABASES
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend Khourie and Crew
/ STREET: One Market Plaza, Steuart St. Tower
/ CITY: San Francisco
```

STATE: CA  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212.433A  
FILING DATE: 14-MAR-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Hughes, Richard L.  
REGISTRATION NUMBER: 31,264  
REFERENCE/DOCKET NUMBER: 16336-2  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-212-433A-33

Query Match 45.5%; Score 25; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
Db 6 QPNY 9

RESULT 6  
US-08-716-256-33  
Sequence 33, Application US/08716256  
Patent No. 6017693  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION  
TITLE OF INVENTION: PATTERNS TO IDENTIFY NUCLEOTIDE, AMINO ACID OR  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: ORGANISMS  
NUMBER OF SEQUENCES: 46  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/716,256  
FILING DATE:  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03239  
FILING DATE: 14-MAR-1995  
APPLICATION NUMBER: US 08/212,433  
FILING DATE: 14-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-2PC  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-716-256-33

Query Match 45.5%; Score 25; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
Db 6 QPNY 9

RESULT 7  
PCT-US95-03239-33  
Sequence 33, Application PC/TUS9503239  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION  
TITLE OF INVENTION: PATTERNS TO IDENTIFY NUCLEOTIDE, AMINO ACID OR  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: ORGANISMS  
NUMBER OF SEQUENCES: 46  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03239  
FILING DATE: 14-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,433  
FILING DATE: 14-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-2PC  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-03239-33

Query Match 45.5%; Score 25; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
Db 6 QPNY 9

RESULT 8  
5436320-3  
Patent No. 5436320  
APPLICANT: SPIEGEL, ALLEN M.  
TITLE OF INVENTION: ANTIBODY REAGENTS THAT IDENTIFY THE  
CARBOXY-TERMINAL PEPTIDE OF THE GTP-BINDING PROTEIN G  
NUMBER OF SEQUENCES: 10  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,377  
FILING DATE: 14-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 564,675

```
/ FILING DATE: 08-AUG-1990
/ APPLICATION NUMBER: 365,919
/ FILING DATE: 15-JUN-1989
/ APPLICATION NUMBER: 100,909
/ FILING DATE: 25-SEP-1987
/ SEQ ID NO:3:
/ LENGTH: 10
5436320-3

Query Match      45.5%; Score 25; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNY 4
        ||||
Db      6 QPNY 9

RESULT 9
5436320-3
; Patent No. 5436320
; APPLICANT: SPIEGEL, ALLEN M.
; TITLE OF INVENTION: ANTIBODY REAGENTS THAT IDENTIFY THE
; CARBOXY-TERMINAL PEPTIDE OF THE GTP-BINDING PROTEIN G
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,377
; FILING DATE: 14-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 564,675
; FILING DATE: 08-AUG-1990
; APPLICATION NUMBER: 365,919
; FILING DATE: 15-JUN-1989
; APPLICATION NUMBER: 100,909
; FILING DATE: 25-SEP-1987
; SEQ ID NO:3:
; LENGTH: 10
5436320-3

Query Match      45.5%; Score 25; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNY 4
        ||||
Db      6 QPNY 9

RESULT 10
US-08-877-605-181
; Sequence 181, Application US/08877605
; Patent No. 6582965
; GENERAL INFORMATION:
; APPLICANT: Robert Townsend
; APPLICANT: Raj Parekh
; APPLICANT: Sally Prime
; APPLICANT: Nick Webb
; TITLE OF INVENTION: A METHOD FOR DE NOVO PEPTIDE SEQUENCE DETERMINATION
; FILE REFERENCE: 9195-004
; CURRENT APPLICATION NUMBER: US/08/877,605
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 181
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide X Library
US-08-877-605-181

Query Match      41.8%; Score 23; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 NYHAV 7
        :|||:
Db      1 DYHAI 5

RESULT 11
US-08-877-605-220
; Sequence 220, Application US/08877605
; Patent No. 6582965
; GENERAL INFORMATION:
; APPLICANT: Robert Townsend
; APPLICANT: Raj Parekh
; APPLICANT: Sally Prime
; APPLICANT: Nick Webb
; TITLE OF INVENTION: A METHOD FOR DE NOVO PEPTIDE SEQUENCE DETERMINATION
; FILE REFERENCE: 9195-004
; CURRENT APPLICATION NUMBER: US/08/877,605
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide X Library
US-08-877-605-220

Query Match      41.8%; Score 23; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      3 NYHAV 7
        :|||:
Db      1 DYHAI 5

RESULT 12
US-09-057-363C-18
; Sequence 18, Application US/09057363C
; Patent No. 6551994
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING THE
; INTERACTION BETWEEN ALPHA-CATENIN AND BETA-CATENIN
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,363C
; FILING DATE: 08-Apr-1998
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 18:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-057-363c-18

Query Match          40.0%; Score 22; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HAVN 8
Db      1 HAVN 4

RESULT 13
US-09-265-107-18
; Sequence 18, Application US/09265107A
; Patent No. 6683048
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR STIMULATING
; FILE REFERENCE: 100086.406C1
; CURRENT APPLICATION NUMBER: US/09/265,107A
; CURRENT FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linear peptide modulating agent
US-09-265-107-18

Query Match          40.0%; Score 22; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HAVN 8
Db      1 HAVN 4

RESULT 14
US-08-092-110A-3
; Sequence 3, Application US/08092110A
; Patent No. 5585477
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND
; METHOD OF INVENTION: METHODS OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/092,110A
; FILING DATE: 13-JUL-1993

Query Match          40.0%; Score 22; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 NYHAVN 8
Db      2 NYHAVN 7

RESULT 15
US-08-273-474-3
; Sequence 3, Application US/08273474
; Patent No. 5691134
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS
; METHOD OF INVENTION: OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,474
; FILING DATE: 13-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-273-474-3

Query Match          40.0%; Score 22; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 NYHAVN 8
Db      2 NYHAVN 7

RESULT 16
US-08-092-110A-3
; Sequence 3, Application US/08273474
; Patent No. 5691134
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS
; METHOD OF INVENTION: OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,474
; FILING DATE: 13-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-273-474-3

Query Match          40.0%; Score 22; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 NYHAVN 8
Db      2 NYHAVN 7
```

Db | | | |  
2 NGHALN 7

Search completed: May 19, 2005, 18:11:53  
Job time : 30.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 18:16:02 ; Search time 38 Seconds  
(without alignments)  
22.788 Million cell updates/sec

Title: US-09-867-159A-5

Perfect score: 57

Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.79.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	45.6	9	2	Ig heavy chain CRD
2	22	38.6	5	2	T-cell receptor be
3	22	38.6	8	2	T-cell receptor ga
4	19	33.3	9	2	adipokinetic hormo
5	18	31.6	6	2	dihydrofolate redu
6	17	29.8	6	2	T-cell receptor be
7	17	29.8	7	2	neuromodulatory pe
8	17	29.8	7	2	neuromodulatory pe
9	17	29.8	7	2	T-cell receptor be
10	16	28.1	6	2	locustakinin - mig
11	16	28.1	6	2	Ig H chain V-D-J r
12	16	28.1	7	2	hypothetical pepti
13	16	28.1	7	4	acylase - Kluyvera
14	16	28.1	8	2	leucokinin V - Mad
15	16	28.1	8	2	leucokinin VII - Ma
16	16	28.1	8	2	leucokinin VII - M
17	16	28.1	9	2	cell surface adhes
18	16	28.1	9	2	dihydrofolate redu
19	15	26.3	6	2	pullulanase (EC 3.
20	15	26.3	7	2	tryptophyllin, bas
21	15	26.3	7	2	adipokinetic hormo
22	15	26.3	8	2	leucokinin VIII -
23	15	26.3	8	2	litorin 2-Glu - Au
24	15	26.3	9	2	litorin I - Austr
25	15	26.3	9	2	Ig heavy chain CDR
26	15	26.3	9	2	gene c-mpl protein
27	15	26.3	9	2	oxytocin-related p
28	15	26.3	9	2	neuropeptide Grb-A
29	15	26.3	9	2	

30 14 24.6 5 2 PT0308 Ig heavy chain CRD  
31 14 24.6 6 2 S66195 alcohol dehydrogen  
32 14 24.6 9 2 A28495 conopressin G - co  
33 14 24.6 9 2 B28495 conopressin S - co  
34 14 24.6 9 2 S07241 litorin - Rohde's  
35 14 24.6 9 2 PT0270 Ig heavy chain CRD  
36 14 24.6 9 2 S39040 lysine-conopressin  
37 13 22.8 5 2 PT0690 T-cell receptor be  
38 13 22.8 6 2 B35640 cerebellar degener  
39 13 22.8 6 2 PT0630 T-cell receptor be  
40 13 22.8 6 2 PD0028 pev-kinin 2 - pena  
41 13 22.8 7 2 S09066 globulin IV alpha  
42 13 22.8 8 2 B44960 neuropeptide Led-C  
43 13 22.8 8 2 S08996 hypertrehalosemic  
44 13 22.8 8 2 B49823 adipokinetic hormo  
45 13 22.8 8 2 A59495 Vesicle associated

#### ALIGNMENTS

##### RESULT 1

PT0299

Ig heavy chain CRD3 region (clone 5-103B) - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C/Accession: PT0299

R/Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A/Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A/Reference number: PT0222; MUID:91108337; PMID:1899102

A/Accession: PT0299

A/Molecule type: DNA

A/Residues: 1-9 <YAM>

A/Experimental source: B lymphocyte

C/Keywords: heterotetramer; immunoglobulin

Query Match 45.6%; Score 26; DB 2; Length 9;

Best Local Similarity 80.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RNSWD 8

Db 1 RESWD 5

##### RESULT 2

PT0580

T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C/Accession: PT0580

R/Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A/Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A/Reference number: PT0509; MUID:91277601; PMID:1711558

A/Accession: PT0580

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-5 <FEE>

A/Experimental source: day 19 fetal thymus, strain BALB/c

C/Keywords: T-cell receptor

Query Match 38.6%; Score 22; DB 2; Length 5;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NSWD 8

Db 2 SSWD 5

##### RESULT 3

A38887  
T-cell receptor gamma chain (St.2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
C:Accession: A38887  
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.  
Mol. Cell. Biol. 11, 5902-5909, 1991  
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene  
A:Reference number: A41946; MUID:92049316; PMID:1658619  
A:Accession: A38887  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-8 <WHE>  
C:Keywords: T-cell receptor

Query Match 38.6%; Score 22; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SWDT 9  
|||  
DB 2 SWDS 5

RESULT 4  
A24244  
adipokinetic hormone - bollworm  
N:Alternate names: Hez-AKH  
C:Species: Heliothis zea (bollworm, corn earworm, tomato fruitworm)  
C:Date: 31-Mar-1988 #sequence\_revision 23-Mar-1995 #text\_change 09-Jul-2004  
C:Accession: A24244  
R:Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Holman, G.M.; Wagner, R.M.; Ridgway  
Biochem. Biophys. Res. Commun. 135, 622-628, 1986  
A:Title: Isolation and primary structure of a peptide from the corpora cardiaca of Helicoverpa zea  
A:Reference number: A24244; MUID:86186794; PMID:3964263  
A:Accession: A24244  
A:Molecule type: protein  
A:Residues: 1-9 <JAF>  
A:Cross-references: UNIPROT:P08901  
C:Superfamily: adipokinetic hormone  
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F;/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 19; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TVRNSW 7  
|||  
DB 3 TFTSSW 8

RESULT 5  
A31263  
dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodium falciparum  
C:Species: Plasmodium falciparum  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
C:Accession: A31263  
R:Peterson, D.S.; Walliker, D.; Welles, T.E.  
Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
A:Reference number: A94217; MUID:89057886; PMID:2904149  
A:Accession: A31263  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-6 <PET>  
C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match 31.6%; Score 18; DB 2; Length 6;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 SWDT 9  
|||  
DB 2 SWES 5

RESULT 6  
PT0519  
T-cell receptor beta chain V-D-J region (100-4C) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0519  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0519  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-6 <FEE>  
A:Experimental source: adult thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 29.8%; Score 17; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WD 8  
|||  
DB 5 WD 6

RESULT 7  
S33244  
neuromodulatory peptide Wamide-1 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S33244  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A:Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of the giant African snail  
A:Reference number: S33244; MUID:93265912; PMID:8495720  
A:Accession: S33244  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>  
A:Cross-references: UNIPROT:P35921

Query Match 29.8%; Score 17; DB 2; Length 7;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
|||  
DB 1 WKEMSVW 7

RESULT 8  
S33245  
neuromodulatory peptide Wamide-2 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S33245  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A:Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of the giant African snail  
A:Reference number: S33244; MUID:93265912; PMID:8495720  
A:Accession: S33245  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>  
A:Cross-references: UNIPROT:P35919

Query Match 29.8%; Score 17; DB 2; Length 7;



Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WTVRNSW 7  
| : |  
1 WREMSVW 7

Db

RESULT 9  
S33246  
neuromodulatory peptide Wamide-3 - giant African snail  
C;Species: Achatina fulica (giant African snail)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S33246  
R;Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A;Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of  
A;Reference number: S33244; MUID:93265912; PMID:8495720  
A;Accession: S33246  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-7 <MIN>  
A;Cross-references: UNIPROT:P35920

Query Match 29.8%; Score 17; DB 2; Length 7;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WTVRNSW 7  
| : |  
1 WKQMSVW 7

Db

RESULT 10  
PT0661  
T-cell receptor beta chain V-D-J region (121-18V) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004  
C;Accession: PT0661  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0661  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-4 <PEE>  
A;Cross-references: UNIPROT:Q8BZQ7; UNIPROT:O8CCN5  
A;Experimental source: day 4 postnatal thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 28.1%; Score 16; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WT 2  
||  
3 WT 4

Db

RESULT 11  
A61068  
locustakinin - migratory locust  
C;Species: Locusta migratoria (migratory locust)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: A61068  
R;Schoofs, L.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.  
Regul. Pept. 37, 49-57, 1992  
A;Title: Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation, pr  
A;Reference number: A61068; MUID:92262851; PMID:1585017  
A;Accession: A61068  
A;Molecule type: protein  
A;Residues: 1-6 <SCH>

A;Cross-references: UNIPROT:P41491  
C;Keywords: amidated carboxyl end; cephalomyotropic peptide; neuropeptide  
P;6/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.1%; Score 16; DB 2; Length 6;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NSW 7  
: ||  
3 SSW 5

Db

RESULT 12  
PH1602  
Ig H chain V-D-J region (wild-type clone 313) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
C;Accession: PH1602  
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.  
J. Exp. Med. 178, 317-329, 1993  
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice  
A;Reference number: PH1580; MUID:93301609; PMID:8315387  
A;Accession: PH1602  
A;Molecule type: DNA  
A;Residues: 1-7 <LEV>  
A;Experimental source: bone marrow pre-B lymphocyte  
C;Keywords: immunoglobulin

Query Match 28.1%; Score 16; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WT 2  
||  
6 WT 7

Db

RESULT 13  
I55382  
hypothetical peptide PA11 promoter region - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 16-Apr-1999 #sequence\_revision 16-Apr-1999 #text\_change 20-Apr-2000  
C;Accession: I55382  
R;Dawson, S.J.; Wiman, B.; Hamsten, A.; Green, F.; Humphries, S.; Henney, A.M.  
J. Biol. Chem. 268, 10739-10745, 1993  
A;Title: The two allele sequences of a common polymorphism in the promoter of the plasmi  
A;Reference number: I55382; MUID:93266509; PMID:8388372  
A;Accession: I55382  
A;Status: translation not shown; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-7 <DAW>  
A;Cross-references: GB:M91557; NID:g190020; PIDN:AAA60110.1; PID:g190021  
C;Comment: This is the hypothetical translation of a sequence from the PA11 gene promote  
C;Genetics:  
A;Gene: GDB:PA11  
A;Cross-references: GDB:120297; OMIM:173360  
A;Map position: 7q21.3-7q22

Query Match 28.1%; Score 16; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WT 2  
||  
1 WT 2

Db

RESULT 14  
S19288  
acylase - Kluyvera cryocrescens  
C;Species: Kluyvera cryocrescens  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C/Accession: S19288  
R/Martin, J.; Slade, A.; Aitken, A.; Arche, R.; Virden, R.  
Biochem. J. 280: 659-662, 1991  
A/Title: Chemical modification of serine at the active site of penicillin acylase from K  
A/Reference number: S19288; MUID:92109664; PMID:1764029  
A/Accession: S19288  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-8 <MAP>  
A/Cross-references: UNIPROT:Q7M124

Query Match 28.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 NSW 7  
|  
Db 2 NMW 4

RESULT 15  
JS0315  
leucokinin V - Madeira cockroach  
C/Species: Leucophaea maderae (Madeira cockroach)  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
C/Accession: JS0315  
R/Holman, G.M.; Cook, B.J.; Nachman, R.J.  
Comp. Biochem. Physiol. C 88, 27-30, 1987  
A/Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic  
A/Reference number: JS0315  
A/Accession: JS0315  
A/Molecule type: protein  
A/Residues: 1-8 <HOL>  
A/Cross-references: UNIPROT:P19987  
C/Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act  
C/Keywords: amidated carboxyl end; cephalomyotropic peptide  
F/8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 NSW 7  
|  
Db 5 SSW 7

Search completed: May 19, 2005, 19:26:29  
Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 18:12:01 ; Search time 173 Seconds  
(without alignments)  
26.640 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1455

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27.5	48.2	9	2 Q38366	Q38366 bacterioph
2	25	43.9	9	2 Q9H3Y3	Q9H3Y3 homo sapien
3	21	36.8	4	1 OCP3 OCTMI	P58649 octopus min
4	21	36.8	7	2 Q8GL04	Q8GL04 borrelia bu
5	21	36.8	8	1 LCK1 LEUMA	P21140 leucophaea
6	21	36.8	8	1 LCK3 LEUMA	P21142 leucophaea
7	21	36.8	8	1 PK1 PERAM	P82685 periplaneta
8	21	36.8	8	1 PK3 PERAM	P82687 periplaneta
9	21	36.8	8	1 PK5 PERAM	P82689 periplaneta
10	21	36.8	8	2 Q8G940	Q8G940 borrelia bu
11	21	36.8	9	2 Q8GL26	Q8GL26 borrelia bu
12	21	36.8	9	2 Q9R9C4	Q9R9C4 borrelia bu
13	19	33.3	8	2 Q37854	Q37854 bacterioph
14	19	33.3	8	2 Q8GL21	Q8GL21 borrelia bu
15	19	33.3	9	1 AKH HELZE	Q67787 heliothis z
16	19	33.3	9	2 Q8GL31	Q8GL31 borrelia bu
17	18	31.6	8	1 C125 CYPDO	P83661 cyphononyx
18	17	29.8	7	1 WWA1_ACHFU	P35919 achatina fu
19	17	29.8	7	1 WWA2_ACHFU	P35920 achatina fu
20	17	29.8	7	1 WWA3_ACHFU	P35921 achatina fu
21	17	29.8	7	2 O49223	O49223 glycine max
22	17	29.8	8	2 Q62721	Q62721 rattus norv
23	17	29.8	8	2 P79940	P79940 xenopus lae
24	17	29.8	9	2 Q6LDB5	Q6LDB5 homo sapien
25	17	29.8	9	2 Q673W5	Q673W5 tyrannus me
26	17	29.8	9	2 Q673W6	Q673W6 terpsiphon
27	17	29.8	9	2 Q673W7	Q673W7 telophorus
28	17	29.8	9	2 Q673W8	Q673W8 tchagra sen
29	17	29.8	9	2 Q673W9	Q673W9 tchagra aus
30	17	29.8	9	2 Q673X0	Q673X0 rhodophoneu
31	17	29.8	9	2 Q673X1	Q673X1 rhipidura a

32	17	29.8	9	2 Q673X2	Q673X2 pseudobias
33	17	29.8	9	2 Q673X3	Q673X3 prionops sc
34	17	29.8	9	2 Q673X4	Q673X4 prionops re
35	17	29.8	9	2 Q673X5	Q673X5 platysteira
36	17	29.8	9	2 Q673X6	Q673X6 oriolus xan
37	17	29.8	9	2 Q673X7	Q673X7 nilaus afer
38	17	29.8	9	2 Q673X8	Q673X8 bias flammu
39	17	29.8	9	2 Q673X9	Q673X9 malaconotus
40	17	29.8	9	2 Q673Y0	Q673Y0 lanius coll
41	17	29.8	9	2 Q673Y1	Q673Y1 lanioturdus
42	17	29.8	9	2 Q673Y2	Q673Y2 laniarius l
43	17	29.8	9	2 Q673Y3	Q673Y3 laniarius f
44	17	29.8	9	2 Q673Y4	Q673Y4 laniarius b
45	17	29.8	9	2 Q673Y5	Q673Y5 laniarius a

ALIGNMENTS

RESULT 1  
Q38366 PRELIMINARY; PRT; 9 AA.  
AC Q38366;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-DEC-2001 (Tremblrel. 01, Last sequence update)  
DE E gene product (Fragment).  
OS Bacteriophage phi-X174.  
OC Viruses; ssDNA viruses; Microviridae; Microvirus.  
OX NCBI\_TaxID=10847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88118956; PubMed=2963134;  
RA Buckley K.J., Hayashi M.;  
RT "Role of premature translational termination in the regulation of  
expression of the phi X174 lysis gene.";  
RL J. Mol. Biol. 198:599-607(1987).  
DR EMBL; X07809; CAA30668.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1207 MW; C093B37731B36412 CRC64;

Query Match 48.2%; Score 27.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.6e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

Qy 1 WTVRNSWDT 9  
||: |||  
Db 4 WTL---WDT 9

RESULT 2  
Q9H3Y3 PRELIMINARY; PRT; 9 AA.  
AC Q9H3Y3  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE DJ839811.1 (Novel protein with a Kunitz/Bovine pancreatic trypsin  
inhibitor domain and WAP-type (Whey Acidic Protein) 'four-disulfide  
core' domains) (Fragment).  
GN Name=dj461p17.1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lloyd D.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL121778; CAB76844.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1036 MW; 2C417B01B412D1B3 CRC64;

```

Query Match      43.9%; Score 25; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVRN 4
DB 2 WTVR 5

RESULT 3
OCP3_OCTMI      STANDARD; PRT; 4 AA.
ID OCP3_OCTMI
AC P58639;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=89766;
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20336815; PubMed=10876044; DOI=10.1016/S0196-9781(00)00201-1;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";
RL Peptides 21:623-630 (2000).
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-4 is a 1000 time less active
CC than Ocp-3.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- FTIM: Ocp-4 has D-Ser instead of L-Ser.
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI; RANGE=1-4; NOTE=Ref.1.
KW D-amino acid; Direct protein sequencing; Hormone.
FT MOD RES 2 2 D-serine (in form Ocp-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match      36.8%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SWD 8
DB 2 SWD 4

RESULT 4
Q8GL04          PRELIMINARY; PRT; 7 AA.
ID Q8GL04;
AC Q8GL04;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PF-50 protein (Fragment).
GN Name=PF-50;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid group cp32-5.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-0107;
RX MEDLINE=22610300; PubMed=12724373; DOI=10.1099/mic.0.26120-0;
RA Miller J.C., Stevenson B.;
RT "Immunological and genetic characterization of Borrelia burgdorferi
RT Rapa and Eppa proteins.";
RL Microbiology 149:1113-1125 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=93-0107;
RX MEDLINE=22990544; PubMed=14629041;

```

---

```

RA Stevenson B., Miller J.C.;
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
RT erp genes generates sequence identity amidst diversity.";
RL J. Mol. Evol. 57:309-324 (2003).
DR EMBL; AY142103; AAN17848.1; -.
KW Plasmid.
FT NON TER 1 1
SQ SEQUENCE 7 AA; 914 MW; 6337244330504310 CRC64;

Query Match      36.8%; Score 21; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5
DB 1 WIKN 5

RESULT 5
LCK1_LEUMA      STANDARD; PRT; 8 AA.
ID LCK1_LEUMA
AC F21140;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Leucokinin I (L-I).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two neuropeptides from
RT Leucophaea maderae: members of a new family of Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:205-211 (1986).
CC -!- FUNCTION: This cephalomyotropic peptide stimulates contractile
CC activity of cockroach prothodum (hindgut).
CC -!- SUBCELLULAR LOCATION: Secreted.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD RES 8 8 Glycine amide.
SQ SEQUENCE 8 AA; 893 MW; DC6365B44CDC76A CRC64;

Query Match      36.8%; Score 21; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NSW 7
DB 5 NSW 7

RESULT 6
LCK3_LEUMA      STANDARD; PRT; 8 AA.
ID LCK3_LEUMA
AC F21142;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Leucokinin III (L-III).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of two additional neuropeptides from
RT Leucophaea maderae: members of a new family of Cephalomyotropins.";

```

RL Comp. Biochem. Physiol. 84C:271-276 (1986).  
CC -1- FUNCTION: This cephalomyotropic peptide stimulates contractile  
CC activity of cockroach prothoracic (hindgut).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 NSW 7  
Db 5 NSW 7  
  
RESULT 7  
PK3\_PERAM STANDARD; PRT; 8 AA.  
AC P82685;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-1 (Pea-K-1).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.  
RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979; DOI=10.1016/S0167-0115(97)01029-X;  
RA Predel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
FT MOD RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 950 MW; 326365B449D5A774 CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 NSW 7  
Db 5 NSW 7  
  
RESULT 8  
PK3\_PERAM STANDARD; PRT; 8 AA.  
AC P82687;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-3 (Pea-K-3).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.

RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979; DOI=10.1016/S0167-0115(97)01029-X;  
RA Predel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
RL Regul. Pept. 71:199-205 (1997).  
CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
CC activity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MASS SPECTROMETRY: MW=907.92; METHOD=Electrospray; RANGE=1-8;  
CC NOTE=Ref.1.  
CC -1- SIMILARITY: Belongs to the kinin family.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 909 MW; DC6365B449D5A76A CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 NSW 7  
Db 5 NSW 7  
  
RESULT 9  
PK5\_PERAM STANDARD; PRT; 8 AA.  
AC P82689;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-5 (Pea-K-5).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.  
RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979; DOI=10.1016/S0167-0115(97)01029-X;  
RA Predel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
RL Regul. Pept. 71:199-205 (1997).  
CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
CC activity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MASS SPECTROMETRY: MW=864.10; METHOD=Electrospray; RANGE=1-8;  
CC NOTE=Ref.1.  
CC -1- SIMILARITY: Belongs to the kinin family.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 865 MW; C76365B449CDC775 CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 NSW 7  
Db 5 NSW 7  
  
RESULT 10  
Q8G940 PRELIMINARY; PRT; 8 AA.  
ID Q8G940  
AC Q8G940;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE PF-50 protein (Fragment).
GN Name=PF-50;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid group cp32-9, and Plasmid group cp32-12.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DNI27C19-2; PLASMID=group cp32-9;
RX MEDLINE=22610300; PubMed=12724373; DOI=10.1099/mic.0.26120-0;
RA Miller J.C., Stevenson B.;
RT "Immunological and genetic characterization of Borrelia burgdorferi
  BapA and EppA proteins.";
RL Microbiology 149:1113-1125 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DNI27C19-2, and Sh-2-82;
RX PLASMID=group cp32-9, and group cp32-12;
RX MEDLINE=22990544; PubMed=14629041;
RA Stevenson B., Miller J.C.;
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
  erp genes generates sequence identity amidst diversity.";
RL J. Mol. Evol. 57:309-324 (2003).
DR EMBL; AY142104; AAN17853.1; -.
DR EMBL; AY142097; AAN17907.1; -.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 8 AA; 1042 MW; 1437244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 8;
Best Local Similarity 40.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5
DB | : |
  2 WIKN 6

RESULT 11
Q8GL26 PRELIMINARY; PRT; 9 AA.
AC Q8GL26;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PF-50 protein (Fragment).
GN Name=PF-50;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid group cp32-5.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sh-2-82;
RX MEDLINE=22990544; PubMed=14629041;
RA Stevenson B., Miller J.C.;
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
  erp genes generates sequence identity amidst diversity.";
RL J. Mol. Evol. 57:309-324 (2003).
DR EMBL; AY142092; AAN17873.1; -.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 9 AA; 1206 MW; 5A4A244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5
DB | : |
  3 WIKN 7
```

```
RESULT 12
Q9R9C4 PRELIMINARY; PRT; 9 AA.
AC Q9R9C4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Borrelia burgdorferi plasmid cp32-2, possible partition proteins,
  complete cds (PF-50 protein) (Fragment).
GN Name=PF-50;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-2, and plasmid group cp32-4.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B31; PLASMID=cp32-2;
RX MEDLINE=98361033; PubMed=9695920;
RA Stevenson B., Casjens S., Rosa P.;
RT "Evidence of past recombination events among the genes encoding the
  Erp antigens of Borrelia burgdorferi.";
RL Microbiology 144:1869-1879 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sh-2-82; PLASMID=group cp32-4;
RX MEDLINE=22990544; PubMed=14629041;
RA Stevenson B., Miller J.C.;
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete
  erp genes generates sequence identity amidst diversity.";
RL J. Mol. Evol. 57:309-324 (2003).
DR EMBL; AF022479; AAC35438.1; -.
DR EMBL; AY142091; AAN17865.1; -.
KW Plasmid.
FT NON_TER
SQ SEQUENCE 9 AA; 1155 MW; 4E1A244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. No. 1.6e+06;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5
DB | : |
  3 WIKN 7

RESULT 13
Q37854 PRELIMINARY; PRT; 8 AA.
AC Q37854;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Coliphage gene of unknown function, 5'end. (Fragment).
OS Bacteriophage R17.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
  Levivirus.
OX NCBI_TaxID=12026;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=73224987; PubMed=4352721;
RA Rensing U.F.E.;
RT "A sequence of seventy-three nucleotides from the coliphage R17
  genome.";
RL Biochem. J. 131:593-604 (1973).
DR EMBL; M24820; AAA72755.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 969 MW; ECB45412C1E72726 CRC64;

Query Match 33.3%; Score 19; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 VRNS 6  
 Db ||||  
 5 VRNS 8

FT MOD RES 9 9 Glycine amide.  
 SQ SEQUENCE 9 AA; 1026 MW; 403665A5A1A9D1A7 CRC64;

Query Match 33.3%; Score 19; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.6e+06;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TVRNSW 7  
 Db |::|  
 3 TTTSSW 8

Search completed: May 19, 2005, 18:25:45  
 Job time : 175 secs

## RESULT 14

Q8GL21 ID\_ Q8GL21 PRELIMINARY; PRT; 8 AA.  
 AC Q8GL21;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE PF-50 protein (Fragment).  
 GN Name=PF-50;  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OG Plasmid group cp32-8.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sh-2-82;  
 RX MEDLINE=22990544; PubMed=14629041;  
 RA Stevenson B., Miller J.C.;  
 RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
 erp genes generates sequence identity amidst diversity.";  
 RL J. Mol. Evol. 57:309-324(2003).  
 DR EMBL; AY142094; AAN17903.1; -.  
 KW Plasmid.  
 FT NON TER 1 1  
 SQ SEQUENCE 8 AA; 1042 MW; 1437244337204373 CRC64;

Query Match 33.3%; Score 19; DB 2; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 1.6e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
 Db |::|  
 2 WILKN 6

## RESULT 15

AKH\_HELZE ID\_AKH\_HELZE STANDARD; PRT; 9 AA.  
 AC P67787; P08901;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Adipokinetic hormone (Hez-AKH).  
 OS Heliothis zea (Corn earworm) (Bollworm).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;  
 OC Noctuidae; Heliothinae; Helicoverpa.  
 OX NCBI\_TaxID=7113;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=86186794; PubMed=3964263;  
 RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Holman G.M.,  
 RA Wagner R.M., Ridgway R.L., Hayes D.K.;  
 RT "Isolation and primary structure of a peptide from the corpora  
 cardiaca of Heliothis zea with adipokinetic activity.";  
 RL Biochem. Biophys. Res. Commun. 135:622-628(1986).  
 CC -1- FUNCTION: This hormone, released from cells in the corpora  
 cardiaca after the beginning of flight, causes release of  
 diglycerides from the fat body and then stimulates the flight  
 muscles to use these diglycerides as an energy source.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the AKH / HRTN / RPCH family.  
 DR PIR; A24244; A24244.  
 DR InterPro; IPR002047; AKH.  
 DR PROSITE; PS00256; AKH; 1.  
 KW Amidation; Direct protein sequencing; Flight; Neuropeptide;  
 KW Pyrrolidone carboxylic acid. Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1

***This Page Blank (uspio)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 18:11:02 ; Search time 160 Seconds  
(without alignments)  
21.755 Million cell updates/sec

Title: US-09-867-159a-5

Perfect score: 57

Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 330156

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	57	100.0	9	5	AAO20571 Cysteine
2	57	100.0	9	5	AB98536 Cysteine
3	41	71.9	8	5	ABJ04065 Cysteine
4	37	64.9	8	2	AAR77352 Cysteine
5	36	63.2	9	4	AAW76056 Trypanos
6	33	57.9	9	2	AAW60399 Tumour ho
7	33	57.9	9	2	AAW93726 Mouse B16
8	33	57.9	9	3	AAW21816 Murine me
9	33	57.9	9	4	AAE06394 Tumour ho
10	30	52.6	5	4	AAE85132 C. parvum
11	30	52.6	5	5	ABJ04066 C parvum
12	30	52.6	5	8	ADQ95813 Cryptospor
13	29	50.9	7	8	ADH12805 Abalone c
14	29	50.9	8	4	AAW85131 P. vincke
15	29	50.9	8	8	ADQ95812 Cryptospor
16	29	50.9	9	8	ADK10416 Human pap
17	29	50.9	9	8	ADK10223 Human pap
18	28	49.1	9	7	ADM18218 Chlamydia
19	28	49.1	9	8	ADR23828 Human CNT
20	27	47.4	6	2	AAW57079 Fasciola
21	27	47.4	9	5	AAU71234 Human MHC
22	26	45.6	7	5	ABG60349 Selective
23	26	45.6	8	7	ADL17323 Human scr
24	26	45.6	9	4	AB66538 Phage clo
25	26	45.6	9	5	ABG97265 Human leu

26	26	45.6	9	5	ABJ16457	Abj16457 Zinc tran
27	26	45.6	9	5	ABJ15476	Abj15476 Zinc tran
28	26	45.6	9	6	ABJ38076	Abj38076 Human cyt
29	26	45.6	9	8	ADN64445	ADN64445 HLA bindi
30	26	45.6	9	8	ADP30382	ADP30382 Human sec
31	26	45.6	9	8	ADR23827	ADR23827 Human CNT
32	25	43.9	7	2	AAW30429	AAW30429 HRE-I aff
33	25	43.9	8	3	AAW77698	AAW77698 TSG-6 bin
34	25	43.9	9	4	AAU24450	AAU24450 Human MHC
35	25	43.9	9	4	AAU24036	AAU24036 Human MHC
36	25	43.9	9	4	AAU23831	AAU23831 Human MHC
37	25	43.9	9	4	AAU24154	AAU24154 Human MHC
38	25	43.9	9	5	ABG34271	ABG34271 Human leu
39	25	43.9	9	7	ADK23765	ADK23765 Human 98P
40	25	43.9	9	8	ADK02690	ADK02690 Hepatitis
41	25	43.9	9	8	ADK07503	ADK07503 Hepatitis
42	24	42.1	7	2	AAW30217	AAW30217 Physarum
43	24	42.1	7	7	ADC54100	ADC54100 Commam am
44	24	42.1	7	7	ADL17066	ADL17066 Phage-dis
45	24	42.1	8	4	ABW78541	ABW78541 HIV-2 gp

## ALIGNMENTS

### RESULT 1

AAO20571  
ID AAO20571 standard; peptide; 9 AA.

XX AAO20571;

XX AAO20571;

DT 02-JAN-2003 (first entry)

XX 02-JAN-2003 (first entry)

DE Cysteine protease epitope peptide region, SEQ ID No 5.

XX Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
XX anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
XX allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
XX atopic eczema; epitope.

OS Dermatophagoides pteronyssinus.

XX WO200278736-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-FR001098.

XX 30-MAR-2001; 2001FR-00004370.

PR 03-MAY-2001; 2001FR-00005929.

PR 29-MAY-2001; 2001US-00867159.

XX (ANTI-) ANTIALIS SARL.

XX Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-750636/81.

XX Antiallergic compositions containing an anti-histamine, a histamine  
XX synthesis inhibitor, and optionally an allergen or nucleic acid coding  
XX for the allergen.

XX Claim 14; Page 11; 32pp; French.

XX The invention relates to antiallergic compositions containing an anti-  
XX histamine, a histamine synthesis inhibitor, and optionally an allergen or  
XX isolated nucleic acid molecule that has at least one polynucleotide  
XX sequence coding for the allergen, together with a pharmaceutical carrier.  
XX The pharmaceutical composition of the invention is useful as a non-  
XX specific antiallergic treatment, and also useful in the treatment of  
XX allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
XX allergic and atopic eczema. This sequence represents a peptide of a  
XX cysteine protease epitope region relating to the antiallergic

CC compositions of the invention  
 XX Sequence 9 AA;  
 SQ

Query Match 100.0%; Score 57; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||||  
 Db 1 WTVRNSWDT 9

RESULT 2  
 ABB98536  
 ID ABB98536 standard; peptide; 9 AA.  
 AC  
 XX ABB98536;  
 XX

DT 13-DEC-2002 (first entry)

DE Cysteine protease epitope #3.

XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunotherapy; allergen; allergic hypersensitivity reaction;  
 KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
 KW cysteine protease.

XX Dermatophagoides pteronyssinus.

XX FR2822709-A1.

PN

PD 04-OCT-2002.

PF 03-MAY-2001; 2001FR-00005929.

PR 30-MAR-2001; 2001FR-00004370.

XX (ANTI-) ANTIALIS SARL.

XX Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-735037/80.

XX Antiallergic composition, useful for preventing and treating e.g. asthma,  
 PT rhinitis or eczema, containing at least two of allergen, antihistamine  
 PT and histamine synthesis inhibitor.

XX Claim 8; Page 6; 33pp; French.

XX The present invention relates to an antiallergic pharmaceutical  
 CC composition (I) comprising a pharmaceutical carrier containing an active  
 CC agent combination of at least two of: an allergen; an antihistamine; and  
 CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
 CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
 CC rhinitis or allergic atopic eczema, in babies, children or adults. The  
 CC present sequence is a peptide fragment (epitope) of cysteine protease  
 CC from Dermatophagoides pteronyssinus, which was used as an allergen in the  
 CC invention

XX Sequence 9 AA;

Query Match 100.0%; Score 57; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||||  
 Db 1 WTVRNSWDT 9

RESULT 3  
 ABJ04065

ID ABJ04065 standard; peptide; 8 AA.  
 XX  
 AC ABJ04065;

DT 27-SEP-2002 (first entry)

DE C parvum cryptopain peptide fragment SEQ ID NO: 112.

XX Cryptosporidium detection; GP900; P68; cryptopain; cryptosporidiosis.

XX Cryptosporidium parvum.

OS WO200194631-A1.

PN 13-DEC-2001.

PD

XX 14-MAY-2001; 2001WO-US015624.

XX 06-JUN-2000; 2000US-00588995.

XX (REGC ) UNIV CALIFORNIA.

XX Petersen C, Barnes DA, Nelson RG, Gut J;

XX WPI; 2002-566447/60.

XX Detecting Cryptosporidium in biological and environmental samples and  
 PT diagnosis of cryptosporidiosis involves, contacting the sample with  
 PT Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or RNA.  
 XX Example 18; Page 154; 157pp; English.

XX The present invention relates to a method of detecting Cryptosporidium in  
 CC biological and environmental samples, and of diagnosing  
 CC cryptosporidiosis. This involves obtaining a sample and contacting it  
 CC with Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or  
 CC RNA, or its variant, mutant or fragment. The method is also useful for  
 CC detecting and identifying individual Cryptosporidium isolates based on  
 CC the genetic characteristics, and for diagnosis of prior or concurrent  
 CC Cryptosporidium infection. The present sequence is a C. parvum peptide  
 CC sequence used in the exemplification of the invention

XX Sequence 8 AA;

Query Match 71.9%; Score 41; DB 5; Length 8;

Best Local Similarity 85.7%; Pred. No. 1.8e+06;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSW 7

Db 2 WTVRNSW 8

RESULT 4

AAR77352

ID AAR77352 standard; peptide; 8 AA.

XX AAR77352;

XX 14-MAY-1996 (first entry)

XX Cysteine proteinase derived peptide #3.

DE Primer; amplify; polymerase chain reaction; PCR; Haemonchus contortus;  
 KW cysteine proteinase; DM.2; DM.4; human; DM.4a; DM.5; antigen;  
 KW vaccine; helminth; parasite; DM.1; ruminant; gastro-intestinal tract.

XX Haemonchus contortus.

XX WO9526402-A1.

XX 05-OCT-1995.

XX

```

PF 24-MAR-1995; 95WO-GB0000665.
XX
XX 25-MAR-1994; 94GB-00005925.
PR 25-MAR-1994; 94GB-00005990.
XX
XX (MLCW ) MALLINCKRODT VETERINARY INC.
XX
XX Knox DP, Smith SK, Smith WD, Redmond D, Murray J;
XX
XX WPI; 1995-351322/45.
DR N-PSDB; AAQ94242.
XX
XX Protective helminth parasite antigen - used in vaccine directed against
PT parasitic nematodes of mammalian gastro-intestinal tract e.g. Haemonchus
PT contortus.
XX
XX Example 16; Fig 15; 79pp; English.
XX
XX The sequences given in AAR77350-53 are peptides derived from the
CC canonical Haemonchus contortus cysteine proteinase molecule which were
CC used in the design of the primers given in AAQ94240-43. These primers
CC were used in the cloning of cDNA fragments from the cysteine proteinase
CC gene, such as DM.1, DM.2, DM.3, DM.4, DM.4a and DM.5 (see also AAQ94246-
CC 51). The amplified fragments may be expressed in a recombinant cell for
CC the production of antigens. These antigens may be used in the preparation
CC of a vaccine against helminth parasites in a human or non-human animal
XX
XX Sequence 8 AA;
SQ
Query Match 64.9%; Score 37; DB 2; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
DB 2 WIVKNSW 8
| : : : : |
| : : : : |

RESULT 5
AAB76056
ID AAB76056 standard; peptide; 9 AA.
XX
XX AAB76056;
AC
XX
XX 10-APR-2001 (first entry)
DT
XX
XX Trypanosoma cruzi HLA-A2 binding peptide.
DE
XX
XX Human leukocyte antigen; HLA; major histocompatibility complex; MHC;
KW cytotoxic T lymphocyte; CTL; human class I MHC; immunogenic;
KW HLA binding peptide; immune response; glycoprotein; cytostatic; virucide;
KW hepatotropic; antiinflammatory; anti-HIV; vaccine;
KW human immunodeficiency virus; proteoacide; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; human papilloma virus; HPV;
KW cytomegalovirus; CMV; acquired immunodeficiency syndrome; AIDS;
KW renal carcinoma; cervical carcinoma; lymphoma; malaria;
KW condyloma acuminatum.
XX
XX Trypanosoma cruzi.
OS
XX
XX WO200100225-A1.
PN
XX
XX 04-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US017842.
PF
XX
XX 29-JUN-1999; 99US-0141422P.
XX
XX (EPIM-) EPIIMUNE INC.
XX
XX Sette A, Sidney J, Southwood S;
PI
XX
XX WPI; 2001-112389/12.
DR

XX Composition comprising human leukocyte antigen binding peptide which
XX comprises isolated, prepared epitope useful for treating viral infections
XX such as acquired immunodeficiency syndrome, and cancer.
XX
XX Claim 1; Page 46; 58pp; English.
XX
XX The present invention describes a composition (I) which comprises at
XX least one human leukocyte antigen (HLA) binding peptide comprising an
XX isolated, prepared epitope comprising one of 547 8-11 residue amino acid
XX sequences (S1), given in AAB75803 to AAB76349. (I) has cytostatic,
XX virucide, hepatotropic, antiinflammatory, anti-HIV (human
XX immunodeficiency virus), and protozoacide activities, which can be used in
XX vaccine production and is an inducer of cytotoxic T-cell response. (I) is
XX useful for inducing a cytotoxic T cell response against a preselected
XX antigen in a patient expressing a specific major histocompatibility
XX complex (MHC) class I allele, by contacting cytotoxic T cells (CTLs) from
XX the patient with (I). (I) is useful as a vaccine to treat and/or prevent
XX viral infection and cancer such as prostate cancer, hepatitis B,
XX hepatitis C, human papilloma virus (HPV) infection, cytomegalovirus
XX (CMV), acquired immunodeficiency syndrome (AIDS), renal carcinoma,
XX cervical carcinoma, lymphoma, malaria, and condyloma acuminatum
XX
XX Sequence 9 AA;
SQ
Query Match 63.2%; Score 36; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
DB 1 WIKNSW 7
| : : : : |
| : : : : |

RESULT 6
AAB60399
ID AAB60399 standard; peptide; 9 AA.
XX
XX AAB60399;
AC
XX
XX 24-AUG-1998 (first entry)
DT
XX
XX Tumour homing peptide of the invention.
DE
XX
XX Tumour homing peptide; in vivo panning; murine melanoma; tumour.
KW
XX
XX Synthetic.
OS
XX
XX WO9810795-A2.
PN
XX
XX 19-MAR-1998.
XX
XX 10-SEP-1997; 97WO-US016086.
PF
XX
XX 10-SEP-1996; 96US-00710067.
PR
XX
XX (BURN-) BURNHAM INST.
PA
XX
XX Ruoslahti E, Pasqualini R;
PI
XX
XX WPI; 1998-207151/18.
DR
XX
XX Tumour homing molecules and their conjugates - useful for, e.g. directing
XX linked moiety to tumour containing angiogenic vasculature.
XX
XX Example 5; Page 80; 105pp; English.
XX
XX AAB60390-432 represent peptides recovered from mouse melanomas. The
XX peptides are tumour homing peptides, and are identified by in vivo
XX panning. The in vivo panning comprises administering a library of diverse
XX peptides to a subject having a tumour, collecting a sample of the tumour,
XX identifying a peptide that homes to the tumour, collecting a sample of
XX normal tissue corresponding to the tumour, and determining that the

```

CC peptide that homes to the tumour is not present in the normal tissue. The  
 CC tumour homing peptides can be linked to a moiety (e.g. doxorubicin), and  
 CC used to direct the moiety to a tumour

XX  
 SQ Sequence 9 AA;

Query Match 57.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 7

AAW93726  
 ID AAW93726 standard; peptide; 9 AA.

XX  
 AC AAW93726;

XX 28-JUN-1999 (first entry)

DE Mouse B16B15b melanoma derived tumour homing peptide 11.

XX Tumour homing peptide; tumour; diagnosis; endothelial cell; melanoma;  
 KW angiogenic vasculature; anti-tumour; anti-inflammatory; anti-angiogenic;  
 KW anti-arthritis; NGR receptor; inhibitor; angiogenesis; anticancer drug;  
 KW prognosis; inflammation; regeneration; wounded tissue; targeting;  
 KW macular degeneration; diabetic retinopathy; rheumatoid arthritis;  
 KW occlusive thrombus; murine.

XX Mus sp.

XX WO9913329-A1.

XX 18-MAR-1999.

XX 08-SEP-1998; 98WO-US018895.

XX 10-SEP-1997; 97US-00926914.

XX 25-AUG-1998; 98US-00139802.

XX (BURN-) BURNHAM INST.

XX Ruoslahti E, Pasqualini R;

XX WPI; 1999-215158/18.

XX Identifying molecules that home to angiogenic vasculature used as targets  
 for anticancer agents.

XX Example VI; Page 115; 180pp; English.

XX This invention describes novel peptides which home to angiogenic  
 CC vasculature, specifically of a tumour and which have anti-tumour, anti-  
 CC inflammatory, anti-angiogenic and anti-arthritis activity. Such molecules  
 CC are identified by treating a purified NGR receptor with a test compound  
 CC and identifying compounds that bind specifically to the NGR receptor. The  
 CC peptides of the invention are inhibitors of angiogenesis and can be used  
 CC to produce conjugates for delivering agents to angiogenic vasculature,  
 CC particularly anticancer drugs or an imaging agent, for diagnosis or  
 CC prognosis. These conjugates may be directed to non-tumour angiogenic  
 CC vasculature, e.g. that present in inflammatory, regenerating or wounded  
 CC tissue, e.g. for treatment of macular degeneration, diabetic retinopathy  
 CC or rheumatoid arthritis. The peptides provide specific targeting to  
 CC tumours, especially their supporting vasculature, since the NGR receptor  
 CC is exposed to the circulation only in angiogenic vasculature. Precise  
 CC targeting should reduce the systemic toxicity of anticancer drugs in the  
 CC conjugates. Complete killing of all target cells may not be essential  
 CC since partial denudation of endothelium may result in an occlusive  
 CC thrombus, and endothelial cells are unlikely to become resistant to  
 CC anticancer agents nor to lose the targeting receptor. AAW93622-W93809 and

CC AAW93843-44 are examples of tumour homing peptides used in the invention  
 XX  
 SQ Sequence 9 AA;

Query Match 57.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 8

AAW93726  
 ID AAW93726 standard; peptide; 9 AA.

XX  
 AC AAW93726;

XX 22-MAR-2001 (first entry)

XX Murine melanoma homing peptide #11.

XX Cytostatic; homing pro-apoptotic conjugate; tumour; antimicrobial;  
 KW breast; prostate; melanoma; cancer; Kaposi's sarcoma; murine.

XX Mus sp.

XX WO200042973-A2.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001602.

XX 22-JAN-1999; 99US-00235902.

XX (BURN-) BURNHAM INST.

XX Ellerby HM, Bredesen DE, Pasqualini R, Ruoslahti EI;

XX WPI; 2000-499174/44.

XX Homing pro-apoptotic conjugate comprising a tumor homing molecule that  
 PT selectively homes to a mammalian cell type or tissue linked to an  
 PT antimicrobial peptide, useful for the treatment of prostate cancer.

XX Example 8; Page 96; 118pp; English.

XX The present invention relates to homing pro-apoptotic conjugates,  
 CC comprising of a tumour homing molecule that selectively homes to a  
 CC mammalian cell type or tissue, linked to an antimicrobial peptide. The  
 CC homing pro-apoptotic conjugates are selectively internalised by the  
 CC mammalian cell type or tissue and exhibits high toxicity, especially to  
 CC angiogenic vasculature. The antimicrobial peptide has low mammalian cell  
 CC toxicity when not linked to the tumor homing molecule. The conjugates are  
 CC useful for the treatment of cancer e.g. Kaposi's sarcoma, breast and  
 CC prostate cancer or melanoma. The present sequence is a homing peptide  
 CC isolated in the present invention, which can be conjugated to an  
 CC antimicrobial peptide to make the homing pro-apoptotic conjugates of the  
 CC present invention

XX Sequence 9 AA;

Query Match 57.9%; Score 33; DB 3; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 9

```

AAE06394
ID  AAE06394 standard; peptide; 9 AA.
XX
AC  AAE06394;
XX
DT  25-SEP-2001 (first entry)
XX
DE  Tumour homing peptide #11 from mouse B16B15b melanoma.
XX
KW  Chimeric prostate-homing pro-apoptotic peptide; prostate-homing peptide;
KW  antimicrobial peptide; prostate cancer; tumour homing molecule; mouse;
KW  cytostatic.
XX
OS  Mus sp.
XX
PN  WO200153342-A1.
XX
PD  26-JUL-2001.
XX
PF  16-JAN-2001; 2001WO-US001362.
XX
PR  21-JAN-2000; 2000US-00489582.
XX
PA  (BURN-) BURNHAM INST.
XX
PI  Ruoslahti EI, Pasqualini R, Arap W, Bredesen DE, Ellerby HM;
XX  WPI; 2001-451901/48.
XX
XX  Novel chimeric prostate-homing pro-apoptotic peptide, used to treat
PT  prostate cancer, comprises a prostate-homing peptide linked to an
PT  antimicrobial peptide.
XX
PS  Example 8; Page 95; 176pp; English.
XX
CC  The patent discloses novel chimeric prostate-homing pro-apoptotic peptide
CC  which comprises a prostate-homing peptide linked to an antimicrobial
CC  peptide, where the chimeric peptide is selectively internalised by and
CC  exhibits high toxicity to prostate tissue and where the antimicrobial
CC  peptide has low mammalian cell toxicity when not linked to prostate-
CC  homing peptide. The chimeric peptide is used to direct an antimicrobial
CC  peptide in vivo to a prostate cancer, to induce selective toxicity in
CC  vivo in a prostate cancer, and to treat a patient with prostate cancer.
CC  The present sequence is a tumour homing peptide from mouse B16B15b
CC  melanoma. This sequence is useful in the homing of pro-apoptotic
CC  conjugates of the invention
XX
SQ  Sequence 9 AA;

Query Match 57.9%; Score 33; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 WTVNSW 7
Db  1 WTCRASW 7

RESULT 10
AAB85132
ID  AAB85132 standard; peptide; 5 AA.
XX
AC  AAB85132;
XX
DT  22-AUG-2001 (first entry)
XX
DE  C. parvum cryptopain protein fragment.
XX
KW  Cryptosporidium; infection; cryptopain; antiparasitic; vaccine;
KW  immunization.
XX
OS  Cryptosporidium parvum.
XX

PN  US6254869-B1.
XX
PD  03-JUL-2001.
XX
PF  27-MAR-1997; 97US-00827171.
XX
PR  27-MAR-1996; 96US-0014233P.
XX
XX  (REGC ) UNIV CALIFORNIA.
XX
PI  Petersen C, Huang J;
XX
XX  WPI; 2001-388923/41.
XX
XX  New vaccines and cryptopain protein for diagnosis and treatment of
PT  Cryptosporidium species infection.
XX
PS  Example 2; Col 19; 32pp; English.
XX
CC  The invention relates to vaccines and cryptopain protein for diagnosis
CC  and treatment of Cryptosporidium species infection. Cryptopain protein
CC  can be expressed by standard recombinant methodology. The vaccines
CC  comprising the cryptopain protein or its fragments are useful for active
CC  immunization of animals and humans against Cryptosporidium infection, or
CC  for production of passive immune products in admixture with an adjuvant.
CC  Cryptopain can be used for prophylactic, therapeutic, diagnostic and
CC  detection purposes. The present sequence represents a peptide fragment of
CC  the C. parvum cryptopain protein
XX
SQ  Sequence 5 AA;

Query Match 52.6%; Score 30; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  3 VRNSW 7
Db  1 VRNSW 5

RESULT 11
ABJ04066
ID  ABJ04066 standard; peptide; 5 AA.
XX
AC  ABJ04066;
XX
DT  27-SEP-2002 (first entry)
XX
DE  C parvum cryptopain peptide fragment SEQ ID NO: 113.
XX
KW  Cryptosporidium detection; GP900; P68; cryptopain; cryptosporidiosis.
XX
OS  Cryptosporidium parvum.
XX
XX  WO200194631-A1.
PN  13-DEC-2001.
XX
XX  14-MAY-2001; 2001WO-US015624.
PF  06-JUN-2000; 2000US-00588995.
PR  (REGC ) UNIV CALIFORNIA.
XX
XX  Petersen C, Barnes DA, Nelson RG, Gut J;
XX  WPI; 2002-566447/60.
XX
XX  Detecting Cryptosporidium in biological and environmental samples and
PT  diagnosis of cryptosporidiosis involves, contacting the sample with
PT  Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or RNA.
XX
PS  Disclosure; Page 154; 157pp; English.

```

XX The present invention relates to a method of detecting *Cryptosporidium* in  
 CC biological and environmental samples, and of diagnosing  
 CC cryptosporidiosis. This involves obtaining a sample and contacting it  
 CC with *Cryptosporidium* GP900, P68 or cryptopain antigen, antibody, DNA or  
 CC RNA, or its variant, mutant or fragment. The method is also useful for  
 CC detecting and identifying individual *Cryptosporidium* isolates based on  
 CC the genetic characteristics, and for diagnosis of prior or concurrent  
 CC *Cryptosporidium* infection. The present sequence is a C. parvum peptide  
 CC sequence used in the exemplification of the invention  
 XX  
 SQ Sequence 5 AA;  
 Query Match 52.6%; Score 30; DB 5; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VRNSW 7  
 Db 1 VRNSW 5  
 |||||  
 RESULT 12  
 ID ADQ95813 standard; peptide; 5 AA.  
 XX  
 AC ADQ95813;  
 XX  
 DT 23-SEP-2004 (first entry)  
 XX  
 DE *Cryptosporidium* parvum cryptopain fragment #3.  
 XX  
 KW *Cryptosporidium* infection; antigen; cryptopain;  
 KW cathepsin L-like cysteine proteinase; passive immunity; vaccine; enzyme.  
 XX  
 OS *Cryptosporidium* parvum.  
 XX  
 PN US6759044-B1.  
 XX  
 PD 06-JUL-2004.  
 XX  
 PF 20-JUN-2000; 2000US-00598062.  
 XX  
 PR 27-MAR-1996; 96US-0014233P.  
 PR 27-MAR-1997; 97US-00827171.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Petersen C, Huang J;  
 DR WPI; 2004-515391/49.  
 XX  
 PS New cryptopain antibodies, useful for treating or inhibiting  
 PT *Cryptosporidium* infection, or as vaccine for to provide passive immunity  
 PT to *Cryptosporidium* infection.  
 XX  
 PS Example 2; SEQ ID NO 15; 33pp; English.  
 XX  
 CC The invention relates to an antibody produced against and binding to  
 CC *Cryptosporidium* antigen, cryptopain. The invention also provides  
 CC polynucleotides encoding cryptopain which is a cathepsin L-like cysteine  
 CC proteinase. The antibody is useful for detecting and treating  
 CC *Cryptosporidium* infection, for providing passive immunity or for  
 CC inhibiting existing *Cryptosporidium* infection. The present sequence is  
 CC *Cryptosporidium* parvum cryptopain active site fragment comprising  
 CC conserved Asn residue.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 52.6%; Score 30; DB 8; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRNSW 7  
 Db 1 VRNSW 5  
 |||||  
 RESULT 13  
 ID ADH12805 standard; peptide; 7 AA.  
 XX  
 AC ADH12805;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Abalone cellulase N-terminal peptide, SEQ ID NO:6.  
 XX  
 KW Abalone; cellulase; liver pancreas; de-inking; decolouring;  
 KW paper manufacture; detergent; food manufacture; animal fodder;  
 KW biomass degradation; energy production; cello-oligosaccharide production;  
 KW cellobiose; cellotriose; waste recycling; EC 3.2.1.4.  
 XX  
 OS Haliotis discus; subsp. hannai.  
 XX  
 PN JP2003235552-A.  
 XX  
 PD 26-AUG-2003.  
 XX  
 PF 13-FEB-2002; 2002JP-00034852.  
 XX  
 PR 13-FEB-2002; 2002JP-00034852.  
 XX  
 PA (HOKK-) HOKKAIDO TLQ KK.  
 DR WPI; 2004-147477/15.  
 XX  
 PT Novel cellulase originating in spiral shells capable of degrading  
 PT cellulose, useful for producing cell oligosaccharide such as cellobiose  
 PT and a celltriose.  
 XX  
 PS Example 3; SEQ ID NO 6; 21pp; Japanese.  
 XX  
 CC The invention relates to a cellulase (ADH12819) from the abalone *Haliotis*  
 CC discus hannai. The cellulase is present in high levels in the liver  
 CC pancreas of the abalone, particularly the anterior portion. The mature  
 CC cellulase (EC 3.2.1.4) has an N-terminal sequence given in ADH12801, and  
 CC has a molecular weight of 66 kD, a pH optimum of 5.5-8.0, an optimum  
 CC temperature of 35-40 degrees Celsius, and is stable at 40 degrees Celsius  
 CC or below. The invention also encompasses a method for the preparation of  
 CC the cellulase from abalone internal organs, and further discloses a cDNA  
 CC sequence (ADH12800) encoding the cellulase. The abalone cellulase is  
 CC useful in the de-inking and decolouring processes in paper manufacturing,  
 CC in detergent compositions, and in the manufacture of foods or animal  
 CC fodder. It is also useful in the degradation of biomass for energy  
 CC production and for the production of cello-oligosaccharides such as  
 CC cellobiose and cellotriose. The cellulase of the invention is extracted  
 CC from parts of abalone which are inedible, thousands of tonnes of which  
 CC are discarded during processing of the shellfish as a foodstuff. The  
 CC method of the invention permits helps to reduce the environmental impact  
 CC of this waste, and permits the inexpensive and rapid production of large  
 CC quantities of cellulase. The present sequence represents an N-terminal  
 CC peptide fragment of abalone cellulase used to design a degenerate PCR  
 CC primer in an example of the invention.  
 XX  
 SQ Sequence 7 AA;  
 Query Match 50.9%; Score 29; DB 8; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TVRNSWD 8  
 Db 1 TISNWD 7  
 ||:|  
 |||

```

RESULT 14
AAB85131
ID AAB85131 standard; peptide; 8 AA.
XX
XX AAB85131;
AC
XX
XX 22-AUG-2001 (first entry)
DT
XX
XX P. vinckei cysteine proteinase conserved motif.
DE
XX
XX Cryptosporidium; infection; cryptopain; antiparasitic; vaccine;
KW immunization; cysteine proteinase.
XX
XX Plasmodium vinckei.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 4
FT /label= Val or Ile
FT
FT Misc-difference 5
FT /label= Lys or Arg
XX
XX US6254869-B1.
PN
XX
XX 03-JUL-2001.
PD
XX
XX 27-MAR-1997; 97US-00827171.
PF
XX
XX 27-MAR-1996; 96US-0014233P.
PR
XX
XX (REGC ) UNIV CALIFORNIA.
PA
XX
XX Petersen C, Huang J;
PI
XX
XX WPI; 2001-388923/41.
DR
XX
XX New vaccines and cryptopain protein for diagnosis and treatment of
PT Cryptosporidium species infection.
XX
XX Example 2; Col 19; 32pp; English.
PS
XX
XX The invention relates to vaccines and cryptopain protein for diagnosis
CC and treatment of Cryptosporidium species infection. Cryptopain protein
CC can be expressed by standard recombinant methodology. The vaccines
CC comprising the cryptopain protein or its fragments are useful for active
CC immunization of animals and humans against Cryptosporidium infection, or
CC for production of passive immune products in admixture with an adjuvant.
CC Cryptopain can be used for prophylactic, therapeutic, diagnostic and
CC detection purposes. The present sequence represents a conserved motif
CC from P. vinckei cysteine proteinase, used for designing degenerate
CC primers for amplifying DNA fragments of C. parvum cryptopain gene
XX
XX Sequence 8 AA;
QY 1 WTVRNSW 7
Db 2 WIXXNSW 8

Query Match 50.9%; Score 29; DB 4; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
Db 2 WIXXNSW 8

Search completed: May 19, 2005, 18:22:51
Job time : 168 secs

RESULT 15
ADQ95812
ID ADQ95812 standard; peptide; 8 AA.
XX
XX ADQ95812;
AC
XX
XX 23-SEP-2004 (first entry)
DT
XX
XX Cryptosporidium parvum cryptopain fragment #2.
DE
XX
XX Cryptosporidium infection; antigen; cryptopain;
KW

```

---

```

KW cathepsin L-like cysteine proteinase; passive immunity; vaccine; enzyme.
XX Cryptosporidium parvum.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 4 /label= Val, Ile
FT
FT Misc-difference 5 /label= Lys, Arg
FT
XX
XX US6759044-B1.
PN
XX
XX 06-JUL-2004.
PD
XX
XX 20-JUN-2000; 2000US-00598062.
PF
XX
XX 27-MAR-1996; 96US-0014233P.
PR
XX
XX 27-MAR-1997; 97US-00827171.
XX
XX (REGC ) UNIV CALIFORNIA.
PA
XX
XX Petersen C, Huang J;
PI
XX
XX WPI; 2004-515391/49.
DR
XX
XX New cryptopain antibodies, useful for treating or inhibiting
PT Cryptosporidium infection, or as vaccine for to provide passive immunity
PT to Cryptosporium infection.
XX
XX Example 2; SEQ ID NO 14; 33pp; English.
PS
XX
XX The invention relates to an antibody produced against and binding to
CC Cryptosporidium antigen, cryptopain. The invention also provides
CC polynucleotides encoding cryptopain which is a cathepsin L-like cysteine
CC proteinase. The antibody is useful for detecting and treating
CC Cryptosporidium infection, for providing passive immunity or for
CC inhibiting existing Cryptosporidium infection. The present sequence is
CC Cryptosporidium parvum cryptopain fragment.
XX
XX Sequence 8 AA;
QY 1 WTVRNSW 7
Db 2 WIXXNSW 8

Query Match 50.9%; Score 29; DB 8; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
Db 2 WIXXNSW 8

Search completed: May 19, 2005, 18:22:51
Job time : 168 secs

```

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 18:25:53 ; Search time 132 Seconds  
(without alignments)  
22.807 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 141056

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	9	10	US-09-867-159A-5
2	33	57.9	9	9	US-09-765-086-116
3	33	57.9	9	14	US-10-264-374-116
4	33	57.9	9	14	US-10-375-992-116
5	33	57.9	9	15	US-10-264-374-116
6	33	57.9	9	16	US-10-375-992-116
7	28	47.1	9	17	US-10-888-348-25
8	27	49.4	9	9	US-09-834-765-155
9	26	45.6	7	16	US-10-363-204-24
10	26	45.6	8	14	US-10-190-082-286
11	26	45.6	9	15	US-10-024-652-130
12	26	45.6	9	15	US-10-024-652-981
13	26	45.6	9	17	US-10-482-284A-134

14	26	45.6	9	17	US-10-888-348-24	Sequence 24, Appl
15	25	43.9	9	10	US-09-793-451-30	Sequence 30, Appl
16	25	43.9	9	10	US-09-793-451-235	Sequence 235, App
17	25	43.9	9	10	US-09-793-451-353	Sequence 353, App
18	25	43.9	9	10	US-09-793-451-649	Sequence 649, App
19	25	43.9	9	10	US-09-942-052-119	Sequence 119, App
20	25	43.9	9	14	US-10-283-722-30	Sequence 30, Appl
21	25	43.9	9	14	US-10-283-722-235	Sequence 235, App
22	25	43.9	9	14	US-10-283-722-353	Sequence 353, App
23	25	43.9	9	14	US-10-283-722-649	Sequence 649, App
24	25	43.9	9	15	US-10-283-903-30	Sequence 30, Appl
25	25	43.9	9	15	US-10-283-903-235	Sequence 235, App
26	25	43.9	9	15	US-10-283-903-353	Sequence 353, App
27	25	43.9	9	15	US-10-283-903-649	Sequence 649, App
28	25	43.9	9	17	US-10-808-187-677	Sequence 677, App
29	24	42.1	7	14	US-10-190-082-16	Sequence 16, Appl
30	24	42.1	8	14	US-10-351-641-1643	Sequence 1643, Ap
31	24	42.1	9	15	US-10-334-726-184	Sequence 184, App
32	24	42.1	9	15	US-10-334-726-295	Sequence 295, App
33	24	42.1	9	15	US-10-609-217-664	Sequence 664, App
34	24	42.1	9	15	US-10-632-388-664	Sequence 664, App
35	24	42.1	9	15	US-10-651-723-664	Sequence 664, App
36	24	42.1	9	15	US-10-645-761-664	Sequence 664, App
37	24	42.1	9	15	US-10-666-696-664	Sequence 664, App
38	24	42.1	9	15	US-10-653-048-664	Sequence 664, App
39	23	40.4	5	16	US-10-705-195-17	Sequence 17, Appl
40	23	40.4	5	16	US-10-714-564A-471	Sequence 471, App
41	23	40.4	6	14	US-10-072-602B-626	Sequence 626, App
42	23	40.4	6	14	US-10-072-602B-628	Sequence 628, App
43	23	40.4	7	14	US-10-072-602B-627	Sequence 627, App
44	23	40.4	7	14	US-10-190-082-129	Sequence 129, App
45	23	40.4	7	14	US-10-190-082-133	Sequence 133, App

ALIGNMENTS

RESULT 1

US-09-867-159A-5  
; Sequence 5, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; TITLE OF INVENTION: and at least one anti-histamine compound  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteromyssinus  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(9)  
; OTHER INFORMATION: Comprises epitope from cystine protease.  
US-09-867-159A-5

Query Match 100.0%; Score 57; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9

Db 1 WTVRNSWDT 9

## RESULT 2

US-09-765-086-116  
; Sequence 116, Application US/09765086  
; Patent No. US20010046498A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Wadhi, Arap  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Ellerby, H. Michael  
; TITLE OF INVENTION: Chimeric Prostate-Homing Peptides With  
; FILE OF INVENTION: Pro-Apoptotic Activity  
; FILE REFERENCE: P-LJ 3844  
; CURRENT FILING DATE: 2001-01-17  
; CURRENT APPLICATION NUMBER: US/09/765,086  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US 09/489,582  
; NUMBER OF SEQ ID NOS: 235  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 116  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-765-086-116

Query Match 57.9%; Score 33; DB 9; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
||| |||  
Db 1 WTCRASW 7

## RESULT 3

US-10-264-374-116  
; Sequence 116, Application US/10264374  
; Publication No. US20030113320A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing  
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using  
; FILE REFERENCE: P-LJ 3203  
; CURRENT APPLICATION NUMBER: US/10/264,374  
; CURRENT FILING DATE: 2002-10-03  
; PRIOR FILING DATE: 2002-10-03  
; PRIOR APPLICATION NUMBER: US/09/139,802  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 08/926,914  
; PRIOR FILING DATE: 1997-09-10  
; PRIOR APPLICATION NUMBER: 08/710,067  
; PRIOR FILING DATE: 1996-09-10  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 116  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-10-264-374-116

Query Match 57.9%; Score 33; DB 14; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
||| |||  
Db 1 WTCRASW 7

## RESULT 4

US-10-375-992-116  
; Sequence 116, Application US/10375992  
; Publication No. US20030152578A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates  
; FILE OF INVENTION: Derived Therefrom, and Methods of Using Same  
; NUMBER OF SEQUENCES: 199  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/375,992  
; FILING DATE: 27-Feb-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/926,914  
; FILING DATE: 10-SEP-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2725  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 116:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; LENGTH: 9 amino acids  
; TOPOLOGY: both  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:  
US-10-375-992-116

Query Match 57.9%; Score 33; DB 14; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
||| |||  
Db 1 WTCRASW 7

## RESULT 5

US-10-264-374-116  
; Sequence 116, Application US/10264374  
; Publication No. US20040096441A9  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing  
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using  
; FILE REFERENCE: P-LJ 3203  
; CURRENT APPLICATION NUMBER: US/10/264,374  
; CURRENT FILING DATE: 2002-10-03  
; PRIOR FILING DATE: 2002-10-03  
; PRIOR APPLICATION NUMBER: US/09/139,802  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 08/926,914  
; PRIOR FILING DATE: 1997-09-10  
US-10-264-374-116

;; PRIOR APPLICATION NUMBER: 08/710,067  
;; PRIOR FILING DATE: 1996-09-10  
;; NUMBER OF SEQ ID NOS: 226  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 116  
;; LENGTH: 9  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
;; OTHER INFORMATION: Peptide  
US-10-264-374-116

Query Match 57.9%; Score 33; DB 15; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
|||  
Db 1 WTCRASW 7

RESULT 6  
US-10-375-992-116  
;; Sequence 116, Application US/10375992  
;; Publication No. US20040131623A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Pasqualini, Renata  
;; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates  
;; DERIVED THEREFROM, AND METHODS OF USING SAME  
;; NUMBER OF SEQUENCES: 199  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell & Flores  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: United States  
;; ZIP: 92122  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/375,992  
;; FILING DATE: 27-Feb-2003  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/926,914  
;; FILING DATE: 10-SEP-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-LJ 2725  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 116:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: both  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 116:  
US-10-375-992-116

Query Match 57.9%; Score 33; DB 16; Length 9;  
Best Local Similarity 71.4%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
|||  
Db 1 WTCRASW 7

us-09-867-159a-5.closed.rapb

Db 1 WTCRASW 7

RESULT 7  
US-10-888-348-25  
;; Sequence 25, Application US/10888348  
;; Publication No. US20050064555A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Marshall, Shannon  
;; APPLICANT: Barbosa, Maria  
;; TITLE OF INVENTION: CILIARY NEUTROPHIC FACTOR VARIANTS  
;; FILE REFERENCE: 34431/US  
;; CURRENT APPLICATION NUMBER: US/10/888,348  
;; CURRENT FILING DATE: 2004-07-09  
;; PRIOR APPLICATION NUMBER: US 60/485,941  
;; PRIOR FILING DATE: 2003-07-09  
;; PRIOR APPLICATION NUMBER: US 60/528,229  
;; PRIOR FILING DATE: 2003-12-08  
;; NUMBER OF SEQ ID NOS: 165  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 25  
;; LENGTH: 9  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-888-348-25

Query Match 49.1%; Score 28; DB 17; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+06;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWD 8  
|||  
Db 1 WTVRSIHD 8

RESULT 8  
US-09-834-765-155  
;; Sequence 155, Application US/09834765  
;; Patent No. US20020055478A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Mary Faria  
;; APPLICANT: Pia M. Challita-Eid  
;; APPLICANT: Arthur B. Raitano  
;; APPLICANT: Steve Chappell Mitchell  
;; APPLICANT: Daniel E.H. Afar  
;; APPLICANT: Ava Jakobovits  
;; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
;; FILE REFERENCE: 129.6US01  
;; CURRENT APPLICATION NUMBER: US/09/834,765  
;; CURRENT FILING DATE: 2001-09-21  
;; PRIOR APPLICATION NUMBER: 60/197,647  
;; PRIOR FILING DATE: 2000-04-12  
;; NUMBER OF SEQ ID NOS: 770  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 155  
;; LENGTH: 9  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-834-765-155

Query Match 47.4%; Score 27; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TVRNSWD 9  
|||  
Db 2 TVLQLMD 9

RESULT 9  
US-10-363-204-24  
;; Sequence 24, Application US/10363204

; Publication No. US20040170955A1  
; GENERAL INFORMATION:  
; APPLICANT: Board of Regents, The University of Texas System  
; TITLE OF INVENTION: Human and Mouse Targeting Peptides Identified by Phage Display  
; FILE REFERENCE: 005774.P003PCT  
; CURRENT APPLICATION NUMBER: US/10/363,204  
; CURRENT FILING DATE: 2003-03-07  
; NUMBER OF SEQ ID NOS: 251  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: (1)..(7)  
; OTHER INFORMATION: synthetic construct  
US-10-363-204-24

Query Match 45.6%; Score 26; DB 16; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVRN 5  
|||:  
Db 2 WTVRD 6

RESULT 10  
US-10-190-082-286  
; Sequence 286, Application US/10190082  
; Publication No. US20030148264A1  
; GENERAL INFORMATION:  
; APPLICANT: Laeky, Lawrence A.  
; APPLICANT: Sidhu, Sachdev S.  
; APPLICANT: Held, Heike A.  
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS  
; FILE REFERENCE: P1905R1  
; CURRENT APPLICATION NUMBER: US/10/190,082  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/303,634  
; PRIOR FILING DATE: 2001-07-06  
; NUMBER OF SEQ ID NOS: 683  
; SEQ ID NO 286  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-190-082-286

Query Match 45.6%; Score 26; DB 14; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TVRNSW 7  
|||:  
Db 2 TVRETW 7

RESULT 11  
US-10-024-652-130  
; Sequence 130, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick

; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; TITLE OF INVENTION: Detection of Cancer  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 130  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-10-024-652-130

Query Match 45.6%; Score 26; DB 15; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TVRNSWDT 9  
|||:  
Db 2 TFR11WDT 9

RESULT 12  
US-10-024-652-981  
; Sequence 981, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Paris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; TITLE OF INVENTION: Detection of Cancer  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 981  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-10-024-652-981

Query Match 45.6%; Score 26; DB 15; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+06;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TVRNSWDT 9  
|||:  
Db 2 TFR11WDT 9

RESULT 13  
US-10-482-284A-134  
; Sequence 134, Application US/10482284A  
; Publication No. US20050019344A1  
; GENERAL INFORMATION:  
; APPLICANT: KHANNA, Rajiv

; APPLICANT: ELKINGTON, Rebecca A.  
; APPLICANT: WALKER, Susan J.  
; TITLE OF INVENTION: Novel human cytomegalovirus (HCMV) cytotoxic T cell epitopes,  
; TITLE OF INVENTION: polypeptides, compositions comprising same and diagnostic and th  
; TITLE OF INVENTION: uses therefore  
; FILE REFERENCE: 47-203  
; CURRENT APPLICATION NUMBER: US/10/482,284A  
; CURRENT FILING DATE: 2003-12-29  
; PRIOR APPLICATION NUMBER: AU PR5931  
; PRIOR FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 318  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 134  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: human cytomegalovirus pp150 CTL epitope peptide  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human cytomegalovirus pp65 C  
; OTHER INFORMATION: epitope peptide  
US-10-482-284A-134

Query Match 45.6%; Score 26; DB 17; Length 9;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRNSWD 8  
|||  
Db 2 VRRSWE 7

RESULT 14  
US-10-888-348-24  
; Sequence 24, Application US/10888348  
; Publication No. US20050064555A1  
; GENERAL INFORMATION:  
; APPLICANT: Marshall, Shannon  
; APPLICANT: Barbosa, Maria  
; TITLE OF INVENTION: CILIARY NEUROTROPHIC FACTOR VARIANTS  
; FILE REFERENCE: 34431/US  
; CURRENT APPLICATION NUMBER: US/10/888,348  
; CURRENT FILING DATE: 2004-07-09  
; PRIOR APPLICATION NUMBER: US 60/485,941  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: US 60/528,229  
; PRIOR FILING DATE: 2003-12-08  
; NUMBER OF SEQ ID NOS: 165  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 24  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-888-348-24

Query Match 45.6%; Score 26; DB 17; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVRN 5  
|||  
Db 4 WTVRS 8

RESULT 15  
US-09-793-451-30  
; Sequence 30, Application US/09793451  
; Publication No. US20030157597A1  
; GENERAL INFORMATION:  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Gazelle S. Rastegar  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Pia M. Challita-Eid

; APPLICANT: Mary Paris  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
; FILE REFERENCE: 129.2USU2  
; CURRENT APPLICATION NUMBER: US/09/793,451  
; CURRENT FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/184,558  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/218,856  
; PRIOR FILING DATE: 2000-07-13  
; NUMBER OF SEQ ID NOS: 752  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-793-451-30

Query Match 43.9%; Score 25; DB 10; Length 9;  
Best Local Similarity 42.9%; Pred. No. 1.3e+06;  
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
|||  
Db 1 WTVSGOW 7

Search completed: May 19, 2005, 18:38:34  
Job time : 133 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: May 19, 2005, 18:17:01 ; Search time 41 Seconds  
(without alignments)  
16.386 Million cell updates/sec

Title: US-09-867-159A-5

Perfect score: 57

Sequence: 1 WTVRNSWD 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 99282

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	71.9	8	4	US-09-588-995A-112
2	33	57.9	9	3	US-09-139-802-116
3	33	57.9	9	4	US-09-659-786-116
4	33	57.9	9	4	US-08-926-914-116
5	30	52.6	5	3	US-08-827-171B-15
6	30	52.6	5	4	US-09-588-995A-113
7	30	52.6	5	4	US-09-588-062-15
8	29	50.9	8	3	US-08-827-171B-14
9	29	50.9	8	4	US-09-588-062-14
10	26.5	46.5	9	2	US-08-706-741B-79
11	26.5	46.5	9	2	US-08-924-695A-79
12	24.5	43.0	9	2	US-08-733-505A-42
13	24.5	43.0	9	2	US-08-706-741B-77
14	24.5	43.0	9	2	US-08-924-695A-77
15	24	42.1	8	3	US-09-082-279B-1488
16	24	42.1	8	3	US-09-315-304B-1642
17	24	42.1	8	4	US-09-834-784-1488
18	24	42.1	8	4	US-09-350-641C-1643
19	24	42.1	9	1	US-08-190-788A-34
20	24	42.1	9	1	US-08-383-474B-39
21	24	42.1	9	1	US-08-465-391A-34
22	24	42.1	9	2	US-08-464-538B-34
23	24	42.1	9	2	US-08-463-076E-78
24	24	42.1	9	4	US-09-428-082B-664
25	23	40.4	5	4	US-09-807-063-17
26	23	40.4	8	3	US-09-082-279B-1483
27	23	40.4	8	3	US-09-082-279B-1484

Sequence 1486, Ap  
Sequence 1489, Ap  
Sequence 1499, Ap  
Sequence 1562, Ap  
Sequence 1637, Ap  
Sequence 1638, Ap  
Sequence 1639, Ap  
Sequence 1640, Ap  
Sequence 1653, Ap  
Sequence 46, Appl  
Sequence 22, Appl  
Sequence 1483, Ap  
Sequence 1484, Ap  
Sequence 1486, Ap  
Sequence 1489, Ap  
Sequence 1499, Ap  
Sequence 1594, Ap  
Sequence 1594, Ap  
Sequence 1562, Ap

## ALIGNMENTS

### RESULT 1

US-09-588-995A-112

; Sequence 112, Application US/09588995A

; Patent No. 6514697

; GENERAL INFORMATION:

; APPLICANT: PETERSEN, CAROLYN

; APPLICANT: BARNES, DEBRA A.

; APPLICANT: NELSON, RICHARD C.

; APPLICANT: GUT, JIRI

; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND

; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM

; TITLE OF INVENTION: INFECTIONS

; FILE REFERENCE: 480.19-5

; CURRENT APPLICATION NUMBER: US/09/588,995A

; CURRENT FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: 08/827,171

; PRIOR FILING DATE: 1997-03-27

; PRIOR APPLICATION NUMBER: 08/928,361

; PRIOR FILING DATE: 1997-09-12

; PRIOR APPLICATION NUMBER: 08/700,651

; PRIOR FILING DATE: 1996-08-14

; PRIOR APPLICATION NUMBER: 08/415,751

; PRIOR FILING DATE: 1995-04-03

; NUMBER OF SEQ ID NOS: 115

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 112

; LENGTH: 8

; TYPE: PRT

; ORGANISM: Cryptosporidium parvum

; FEATURE:

; NAME/KEY: MOD\_RES

; LOCATION: (3)\_

; OTHER INFORMATION: L or I

; US-09-588-995A-112

Query Match 71.9%; Score 41; DB 4; Length 8;

Best Local Similarity 85.7%; Pred. No. 4.1e+05;

Mismatches 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 WTVRNSW 7

|||||

Db

2 WTVRNSW 8

### RESULT 2

US-09-139-802-116

; Sequence 116, Application US/09139802

; Patent No. 6180084

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

```
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/139,802
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926,914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710,067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-139-802-116

Query Match          57.9%; Score 33; DB 3; Length 9;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
Db      1 WTCRASW 7

RESULT 3
US-09-659-786-116
; Sequence 116, Application US/09659786
; Patent No. 6491894
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/659,786
; EARLIER FILING DATE: 2000-09-11
; EARLIER APPLICATION NUMBER: 08/926,914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710,067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-659-786-116

Query Match          57.9%; Score 33; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
Db      1 WTCRASW 7

RESULT 4
US-08-926-914-116
; Sequence 116, Application US/08926914
; Patent No. 6576239
; GENERAL INFORMATION:
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates
; TITLE OF INVENTION: Derived Therefrom, and Methods of Using Same
; NUMBER OF SEQUENCES: 199
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,914
; FILING DATE: 10-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-926-914-116

Query Match          57.9%; Score 33; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
Db      1 WTCRASW 7

RESULT 5
US-08-827-171B-15
; Sequence 15, Application US/08827171B
; Patent No. 6254869
; GENERAL INFORMATION:
; APPLICANT: CAROLYN PETERSEN
; APPLICANT: JIN-XING HUANG
; TITLE OF INVENTION: CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
; TITLE OF INVENTION: PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
; TITLE OF INVENTION: TREATMENT, DIAGNOSIS AND
; TITLE OF INVENTION: DETECTION OF
; TITLE OF INVENTION: CRYPTOSPORIDIUM PARVUM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIK A
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: Wordperfect 6.0a WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,171B
; FILING DATE:
```

```
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates
; TITLE OF INVENTION: Derived Therefrom, and Methods of Using Same
; NUMBER OF SEQUENCES: 199
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,914
; FILING DATE: 10-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-926-914-116

Query Match          57.9%; Score 33; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
Db      1 WTCRASW 7

RESULT 5
US-08-827-171B-15
; Sequence 15, Application US/08827171B
; Patent No. 6254869
; GENERAL INFORMATION:
; APPLICANT: CAROLYN PETERSEN
; APPLICANT: JIN-XING HUANG
; TITLE OF INVENTION: CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
; TITLE OF INVENTION: PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
; TITLE OF INVENTION: TREATMENT, DIAGNOSIS AND
; TITLE OF INVENTION: DETECTION OF
; TITLE OF INVENTION: CRYPTOSPORIDIUM PARVUM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIK A
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: Wordperfect 6.0a WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,171B
; FILING DATE:
```



```
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,233
; FILING DATE: March 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana Verny
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-827-171B-15

Query Match 52.6%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VRNSW 7
Db 1 VRNSW 5

RESULT 6
US-09-588-995A-113
; Sequence 113, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
; US-09-588-995A-113

Query Match 52.6%; Score 30; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VRNSW 7
Db 1 VRNSW 5

RESULT 7
US-09-598-062-15
; Sequence 15, Application US/09598062
; Patent No. 6759044
```

```
; GENERAL INFORMATION:
; APPLICANT: CAROLYN PETERSEN
; APPLICANT: JIN-XING HUANG
; TITLE OF INVENTION: CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
; PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
; TREATMENT, DIAGNOSIS AND
; DETECTION OF
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIK A
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
; COMPUTER: PC
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: Wordperfect 6.0a WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,062
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,171
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/014,233
; FILING DATE: March 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana Verny
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
; US-09-598-062-15

Query Match 52.6%; Score 30; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VRNSW 7
Db 1 VRNSW 5

RESULT 8
US-08-827-171B-14
; Sequence 14, Application US/08827171B
; Patent No. 6254869
; GENERAL INFORMATION:
; APPLICANT: CAROLYN PETERSEN
; APPLICANT: JIN-XING HUANG
; TITLE OF INVENTION: CRYPTOPAIN VACCINES, ANTIBODIES, PROTEINS,
; PEPTIDES, DNA AND RNAs FOR PROPHYLAXIS,
; TREATMENT, DIAGNOSIS AND
; DETECTION OF
; TITLE OF INVENTION: CRYPTOSPORIDIUM PARVUM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIK A
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: California
```

```

; COUNTRY: United States of America
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 KB storage
; COMPUTER: PC
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: Wordperfect 6.0a WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,171B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,233
; FILING DATE: March 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana VERNY
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptides
; FEATURE:
; NAME/KEY: Xaa at 4 is Val/Ile
; NAME/KEY: Xaa at 5 is Lys/Arg
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
; US-08-827-171B-14

```

```

; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/014,233
; FILING DATE: March 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hana Verry
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-1677
; TELEFAX: (415) 324-1678
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-598-062-14

Query Match          50.9%; Score 29; DB 4; Length 8;
Best Local Similarity 57.1%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 3; Indels

Qy 1 WTVRNSW 7
   | |||
Db 2 WIXXNSW 8

RESULT 10
US-08-706-741b-79
; Sequence 79, Application US/08706741B
; Patent No. 595593
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,741B
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-706-741b-79

Query Match          46.5%; Score 26.5; DB 2; Length 9;
Best Local Similarity 55.6%; Pred. No. 4.1e+05;
Matches 5; Conservative 1; Mismatches 2; Indels

Qy 1 WT-VRNSWD 8
   | | | |
Db 1 WTRIIQSWD 9

```

RESULT 11  
US-08-924-695A-79  
; Sequence 79, Application US/08924695A  
; Patent No. 5998583  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/924,695A  
; FILING DATE: 09-SEP-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971798  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-924-695A-79

Query Match 46.5%; Score 26.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 4.1e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 WT-VNRWSD 8  
|| : |||  
Db 1 WTRIQSWD 9

RESULT 12  
US-08-733-505A-42  
; Sequence 42, Application US/08733505A  
; Patent No. 5856445  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: SERINE SUBSTITUTED MUTANTS OF  
; TITLE OF INVENTION: BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/733,505A

; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965458  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-733-505A-42

Query Match 43.0%; Score 24.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 WTVRN-SWD 8  
| : |||  
Db 1 WKEHNRWD 9

RESULT 13  
US-08-706-741B-77  
; Sequence 77, Application US/08706741B  
; Patent No. 5955593  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63146  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/706,741B  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 77:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-706-741B-77

Query Match 43.0%; Score 24.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 WTVRN-SWD 8  
| : |||

Db 1 WKEHNRSD 9

## RESULT 14

US-08-924-695A-77  
; Sequence 77, Application US/08924695A  
; Patent No. 5998583  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/924,695A  
; FILING DATE: 09-SEP-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971798  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 77:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-924-695A-77

Query Match 43.0%; Score 24.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 4.1e+05;  
Matches 5; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 WTVRN-SWD 8  
Db 1 WKEHNRSD 9

## RESULT 15

US-09-082-279B-1488  
; Sequence 1488, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohmed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1488  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: HIV-2

US-09-082-279B-1488

Query Match 42.1%; Score 24; DB 3; Length 8;  
Best Local Similarity 42.9%; Pred. No. 4.1e+05;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
Db 1 WDVFSNW 7

Search completed: May 19, 2005, 18:27:15  
Job time : 42 secs

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

This Page Blank (uspto)